Globally Hyperbolic Regularization of Grad's Moment System

Zhenning Cai*, Yuwei Fan†, Ruo Li‡

February 24, 2013

Abstract

In this paper, we propose a globally hyperbolic regularization to the general Grad's moment system in multi-dimensional spaces. Systems with moments up to an arbitrary order are studied. The characteristic speeds of the regularized moment system can be analytically given and only depend on the macroscopic velocity and the temperature. The structure of the eigenvalues and eigenvectors of the coefficient matrix is fully clarified. The regularization together with the properties of the resulting moment systems is consistent with the simple one-dimensional case discussed in [1]. Besides, all characteristic waves are proven to be genuinely nonlinear or linearly degenerate, and the studies on the properties of rarefaction waves, contact discontinuities and shock waves are included.

Keywords: Grad's moment system; regularization; global hyperbolicity; characteristic wave

1 Introduction

The kinetic gas theory, which is based on the Boltzmann equation, is one of the fundamental tools in modelling non-equilibrium processes. Nevertheless, in most cases, a direct numerical discretization of the Boltzmann equation leads to unacceptable computational costs. In 1940s, Grad [7] proposed the moment approximation of the distribution function, trying to establish a series of intermediate models between the fluid dynamics and the kinetic theory. However, due to a number of defects in Grad's 13-moment equations, such as the appearance of unphysical subshocks, nonexistence of an entropy function, and lack of global hyperbolicity, not much attention is paid to the moment method in the last century.

In the recent twenty years, as the investigation into the moment method becomes deeper, various "regularizations" are proposed to challenge the traditional accusations on the moment method. A list of relevant publications can be found in the references of [13]. Recently, we are interested in the large moment system together with its numerical methods [2, 5, 4, 3], and it is found that the lack of the well-posedness due to the loss of global hyperbolicity is a major obstacle in our simulations, especially for large Mach

*School of Mathematical Sciences, Peking University, Beijing, China, email: `caizn@pku.edu.cn`.

†School of Mathematical Sciences, Peking University, Beijing, China, email: `ywfan@pku.edu.cn`.

‡CAPT, LMAM & School of Mathematical Sciences, Peking University, Beijing, China, email: `rli@math.pku.edu.cn`.

arXiv:1203.0376v1 [math-ph] 2 Mar 2012

number gas flows [5]. Torrilhon [14] provided a 13-moment hyperbolic moment system based on multi-variate Pearson-IV distributions, but it seems unlikely to extend the same technique to systems with large number of moments. As discussed in [13], Levermore [9] gave a partial answer to the question of hyperbolicity of large moment system based on a maximal entropy distribution function. However, the analytical forms of Levermore's equations cannot be obtained once the number of moments is greater than 10. While exploring the method ensuring the hyperbolicity of the moment system, we discovered [1] that the structure of the characteristic polynomial of Grad's moment equations with one-dimensional microscopic velocity is rather simple; thus a globally hyperbolic regularization can be achieved by simply adding two terms to the equation of the highest order moment.

In this paper, the results in [1] are extended to the multi-dimensional space. For multi-dimensional moment systems, the regularization method is consistent with the one-dimensional case. Due to the complexity of the moment systems, this paper is mainly devoted to a rigorous proof of the hyperbolicity of regularized moment system for any space dimensions and an arbitrary order of moments. The result is obtained by firstly restricting the spatial variable in the one-dimensional space, and then it is generalized to the multi-dimensional space using the rotation invariance of the regularized system. For the case of one-dimensional spatial variable, the structure of the coefficient matrix is similar as the Hessenberg matrix, which enables us to calculate the eigenvectors for a given eigenvalue. Then, the hyperbolicity of the moment system follows by counting the number of linearly independent eigenvectors. At the same time, the expressions of all characteristic speeds are obtained, each of which is a sum of the macroscopic velocity and the square root of the temperature scaled by a zero of the Hermite polynomial. Besides, we prove that each characteristic filed of the hyperbolic moment systems is either genuinely nonlinear or linearly degenerate, and some properties of the rarefaction waves, the contact discontinuities and the shock waves are investigated.

The rest of this paper is arranged as follows: in Section 2, a brief review on the moment methods of Boltzmann equation and the results in [1] are presented. Section 3 gives the globally hyperbolic regularization for moment system with one-dimensional spatial variable and multi-dimensional microscopic velocities. And in Section 4, the result for full multi-dimensional moment system is proved. The study on the characteristic waves is carried out in Section 5.

2 Preliminaries

In this section, a concise introduction of the Boltzmann equation is presented. And then some results of the work on the moment method in [5, 1] are briefly reviewed.

2.1 Moment methods for Boltzmann equation

Let the motion of particles be depicted by the distribution function $f(t, \boldsymbol{x}, \boldsymbol{\xi})$ governed by the Boltzmann transport equation

$$\frac{\partial f}{\partial t} + \sum_{j=1}^{D} \xi_j \frac{\partial f}{\partial x_j} = Q(f, f), \qquad t \in \mathbb{R}^+, \quad \boldsymbol{x}, \boldsymbol{\xi} \in \mathbb{R}^D, \tag{2.1}$$

where t denotes the time, $\boldsymbol{x} = (x_1, \cdots, x_D)$ and $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_D)$ stand for the spatial coordinates and the microscopic velocity, respectively. The right hand side $Q(f, f)$ is the

collision term describing the interaction between particles. In this paper, we are focusing on the transportation part, thus the collisionless Boltzmann equation with vanished $Q(f, f)$ is considered.

The moment method proposed by Grad [7] approximates the distribution function by a finite set of moments. To achieve this, we expand f into the Hermite series as in [2]:

$$f(t, \boldsymbol{x}, \boldsymbol{\xi}) = \sum_{\alpha \in \mathbb{N}^D} f_\alpha(t, \boldsymbol{x}) \mathcal{H}_{\theta(t, \boldsymbol{x}), \alpha} \left(\frac{\boldsymbol{\xi} - \boldsymbol{u}(t, \boldsymbol{x})}{\sqrt{\theta(t, \boldsymbol{x})}} \right), \tag{2.2}$$

where $\alpha = (\alpha_1, \cdots, \alpha_D)$ is a D-dimensional multi-index, and the basis functions are defined as

$$\mathcal{H}_{\theta, \alpha}(\boldsymbol{z}) = \prod_{d=1}^{D} \frac{1}{\sqrt{2\pi}} \theta^{-\frac{\alpha_d+1}{2}} He_{\alpha_d}(z_d) \exp\left(-\frac{z_d^2}{2}\right), \quad \boldsymbol{z} = (z_1, \cdots, z_D) \in \mathbb{R}^D, \tag{2.3}$$

where He_k is the k-th degree Hermite polynomial:

$$He_k(x) = (-1)^k \exp\left(\frac{x^2}{2}\right) \frac{\mathrm{d}^k}{\mathrm{d}x^k} \exp\left(-\frac{x^2}{2}\right), \quad k \in \mathbb{N}. \tag{2.4}$$

In (2.2), $\boldsymbol{u}(t, \boldsymbol{x}) = (u_1(t, \boldsymbol{x}), \cdots, u_D(t, \boldsymbol{x}))$ and $\theta(t, \boldsymbol{x})$ denote the macroscopic velocity and temperature, respectively, and they are related to f by

$$\rho(t, \boldsymbol{x}) = \int_{\mathbb{R}^D} f(t, \boldsymbol{x}, \boldsymbol{\xi}) \, \mathrm{d}\boldsymbol{\xi},$$

$$\rho(t, \boldsymbol{x}) \boldsymbol{u}(t, \boldsymbol{x}) = \int_{\mathbb{R}^D} \boldsymbol{\xi} f(t, \boldsymbol{x}, \boldsymbol{\xi}) \, \mathrm{d}\boldsymbol{\xi}, \tag{2.5}$$

$$\rho(t, \boldsymbol{x}) |\boldsymbol{u}(t, \boldsymbol{x})|^2 + D\rho(t, \boldsymbol{x}) \theta(t, \boldsymbol{x}) = \int_{\mathbb{R}^D} |\boldsymbol{\xi}|^2 f(t, \boldsymbol{x}, \boldsymbol{\xi}) \, \mathrm{d}\boldsymbol{\xi},$$

where ρ stands for the density of the gas. The following relations can be deduced from the orthogonality of Hermite polynomials:

$$f_0 = \rho, \quad f_{e_j} = 0, \quad \sum_{d=1}^{D} f_{2e_d} = 0, \quad j = 1, \cdots, D, \tag{2.6}$$

where e_j is the D-dimensional multi-index with its j-th component to be the only nonzero one and equals to 1.

The moment system has been deduced in [5], and here we directly present the result therein:

$$\left(\frac{\partial f_\alpha}{\partial t} + \sum_{d=1}^{D} \frac{\partial u_d}{\partial t} f_{\alpha-e_d} + \frac{1}{2} \frac{\partial \theta}{\partial t} \sum_{d=1}^{D} f_{\alpha-2e_d} \right)$$

$$+ \sum_{j=1}^{D} \left(\theta \frac{\partial f_{\alpha-e_j}}{\partial x_j} + u_j \frac{\partial f_\alpha}{\partial x_j} + (\alpha_j + 1) \frac{\partial f_{\alpha+e_j}}{\partial x_j} \right)$$

$$+ \sum_{j=1}^{D} \sum_{d=1}^{D} \frac{\partial u_d}{\partial x_j} \left(\theta f_{\alpha-e_d-e_j} + u_j f_{\alpha-e_d} + (\alpha_j + 1) f_{\alpha-e_d+e_j} \right)$$

$$+ \frac{1}{2} \sum_{j=1}^{D} \sum_{d=1}^{D} \frac{\partial \theta}{\partial x_j} \left(\theta f_{\alpha-2e_d-e_j} + u_j f_{\alpha-2e_d} + (\alpha_j + 1) f_{\alpha-2e_d+e_j} \right) = 0, \quad \forall \alpha \in \mathbb{N}^D. \tag{2.7}$$

3

In this equation, f_β is taken as zero if any components of β is negative. Some special choices of α lead to the classic hydrodynamic equations:

$$\frac{\partial \rho}{\partial t} + \sum_{j=1}^{D} \left(u_j \frac{\partial \rho}{\partial x_j} + \rho \frac{\partial u_j}{\partial x_j} \right) = 0, \tag{2.8a}$$

$$\rho \frac{\partial u_i}{\partial t} + \sum_{j=1}^{D} \left(\rho u_j \frac{\partial u_i}{\partial x_j} + \frac{\partial p_{e_i+e_j}}{\partial x_j} \right) = 0, \quad i = 1, \cdots, D, \tag{2.8b}$$

$$\frac{D}{2} \rho \frac{\partial \theta}{\partial t} + \sum_{j=1}^{D} \left(\frac{D}{2} \rho u_j \frac{\partial \theta}{\partial x_j} + \frac{\partial q_j}{\partial x_j} \right) + \sum_{i=1}^{D} \sum_{j=1}^{D} p_{e_i+e_j} \frac{\partial u_i}{\partial x_j} = 0, \tag{2.8c}$$

where $p_{e_i+e_j}$ is the pressure tensor[1] and q_j is the heat flux. They are defined as

$$p_{e_i+e_j} = \int_{\mathbb{R}^D} (\xi_i - u_i)(\xi_j - u_j) f \, \mathrm{d}\boldsymbol{\xi} = \delta_{ij} \rho \theta + (1 + \delta_{ij}) f_{e_i+e_j}, \tag{2.9a}$$

$$q_j = \frac{1}{2} \int_{\mathbb{R}} |\boldsymbol{\xi} - \boldsymbol{u}|^2 (\xi_j - u_j) f \, \mathrm{d}\boldsymbol{\xi} = 2 f_{3e_j} + \sum_{d=1}^{D} f_{e_j + 2e_d}, \tag{2.9b}$$

where δ is Kronecker's delta symbol. We refer the readers to [5] for the detailed derivation of (2.8).

Since (2.7) forms an infinite set of moment equations which are not suitable for practical use, the moment closure is in need. The simplest way is to select an integer $M \geqslant 3$ and force $f_\alpha = 0$ if $|\alpha| > M$, and the result is the Grad-type system with $\binom{M+D}{D}$ moments.

2.2 Regularization with 1D velocity space

It is well known that the lack of global hyperbolicity is one of the major defects of Grad's moment equations. For the thirteen moment case, the hyperbolicity region has been analytically obtained in [11]. The construction of globally hyperbolic moment systems is very meaningful to the robustness of fluid simulation using moment approximation. In this direction, a general method by Levermore in [9] on the construction of symmetric hyperbolic moment systems is proposed. Later, Torillhon [14] raises a clever idea to enlarge the hyperbolicity region of the 13-moment system by using Pearson-IV-distributions. In [1], we have studied the general 1D moment systems and found a way to make globally hyperbolic regularization based on the characteristic speed correction. Here we are going to give a brief review on the results therein.

When $D = 1$, the multi-index α becomes a natural number. Substituting (2.8b) and (2.8c) into (2.7), we can eliminate the time derivative of the velocity and temperature. Thus the M-th order Grad's moment system can be written in the form of a quasi-linear system

$$\frac{\partial \boldsymbol{w}}{\partial t} + \mathbf{A}(\boldsymbol{w}) \frac{\partial \boldsymbol{w}}{\partial x} = 0, \tag{2.10}$$

where $\mathbf{A}$ is a matrix dependent on $\boldsymbol{w}$, and

$$\boldsymbol{w} = (\rho, u, \theta, f_3, \cdots, f_M). \tag{2.11}$$

In [1], we have obtained the following results:

---

[1]In some literatures, the pressure tensor is denoted as p_{ij}, $i, j = 1, \cdots, D$. Here the special subscript $e_i + e_j$ is used to match the form of general moments $f_{e_i+e_j}$ for convenience in later use.

1. The characteristic polynomial of $\mathbf{A}(\boldsymbol{w})$ is

$$
\begin{aligned}
|\lambda \mathbf{I} - \mathbf{A}| &= \theta^{\frac{M+1}{2}} He_{M+1}\left(\frac{\lambda - u}{\sqrt{\theta}}\right) \\
&\quad - \frac{(M+1)!}{2\rho}\left[\left((\lambda - u)^2 - \theta\right)f_{M-1} + 2(\lambda - u)f_M\right].
\end{aligned}
\tag{2.12}
$$

2. By adding the regularization term based on characteristic speed correction

$$
\mathcal{R}_M = \frac{M+1}{2}\left(2f_M\frac{\partial u}{\partial x} + f_{M-1}\frac{\partial \theta}{\partial x}\right)
\tag{2.13}
$$

to the right hand side of the last equation of (2.10), the system is turned to be globally hyperbolic and the eigenvalues of the regularized moment system are

$$
u + \mathrm{C}_{j,M+1}\sqrt{\theta}, \quad j = 1, \cdots, M+1,
\tag{2.14}
$$

where $\mathrm{C}_{j,k}$ is the j-th root of the Hermite polynomial $He_k(x)$, noticing that $He_k(x), k \in \mathbb{N}$ has k different zeros, which read $\mathrm{C}_{1,k}, \cdots, \mathrm{C}_{k,k}$ and satisfy $\mathrm{C}_{1,k} < \cdots < \mathrm{C}_{k,k}$.

The second result gives a practical implementation of a globally hyperbolic regularization.

2.3 Reformulation of the moment system

In order to facilitate the studying of the moment system when $D \geq 2$, we rewrite (2.7) in another form. Let $p = \rho\theta$, then $p = \frac{1}{D}\sum_{d=1}^{D} p_{2e_d}$, and we have

$$
\frac{\partial \theta}{\partial x_j} = -\frac{\theta}{\rho}\frac{\partial \rho}{\partial x_j} + \frac{1}{D\rho}\sum_{d=1}^{D}\frac{\partial p_{2e_d}}{\partial x_j}, \quad j = 1, \cdots, D.
\tag{2.15}
$$

By substituting (2.8) and (2.15) into (2.7), the following equation is obtained with some simplification:

$$
\begin{aligned}
&\frac{\partial f_\alpha}{\partial t} + \sum_{j=1}^{D}\left(\theta\frac{\partial f_{\alpha-e_j}}{\partial x_j} + u_j\frac{\partial f_\alpha}{\partial x_j} + (\alpha_j+1)\frac{\partial f_{\alpha+e_j}}{\partial x_j}\right) + \sum_{j=1}^{D}\left(-\frac{\theta}{2\rho}C_{\theta,\alpha}^{(j)}\right)\frac{\partial \rho}{\partial x_j} \\
&+ \sum_{j=1}^{D}\sum_{d=1}^{D}\frac{\partial u_d}{\partial x_j}\left(\theta f_{\alpha-e_d-e_j} + (\alpha_j+1)f_{\alpha-e_d+e_j} - \frac{C_\alpha}{D\rho}p_{e_j+e_d}\right) \\
&+ \sum_{j=1}^{D}\sum_{d=1}^{D}\left(\left(-\frac{f_{\alpha-e_d}}{\rho}\right)\frac{\partial p_{e_j+e_d}}{\partial x_j} + \frac{C_{\theta,\alpha}^{(j)}}{2D\rho}\frac{\partial p_{2e_d}}{\partial x_j}\right) + \left(-\frac{C_\alpha}{D\rho}\right)\sum_{j=1}^{D}\frac{\partial q_j}{\partial x_j} = 0,
\end{aligned}
\tag{2.16}
$$

where C_α and $C_{\theta,\alpha}^{(j)}$ are defined as

$$
C_\alpha = \sum_{k=1}^{D} f_{\alpha-2e_k},
\tag{2.17a}
$$

$$
C_{\theta,\alpha}^{(j)} = \sum_{k=1}^{D}\left(\theta f_{\alpha-2e_k-e_j} + (\alpha_j+1)f_{\alpha-2e_k+e_j}\right).
\tag{2.17b}
$$

Then collecting (2.8), (2.16) and (2.9a), we get

$$
\frac{\partial p_{2e_i}/2}{\partial t} + \sum_{j=1}^{D} u_j \frac{\partial p_{2e_i}/2}{\partial x_j} + \sum_{j=1}^{D}\left(\frac{1}{2}+\delta_{ij}\right)\rho\theta\frac{\partial u_j}{\partial x_j}
$$
$$
+ \sum_{j=1}^{D}\sum_{d=1}^{D}(2\delta_{ij}+1)f_{2e_i-e_d+e_j}\frac{\partial u_d}{\partial x_j} + \sum_{j=1}^{D}(2\delta_{ij}+1)\frac{\partial f_{2e_i+e_j}}{\partial x_j} = 0, \quad i = 1, \cdots, D.
\tag{2.18}
$$

The (2.8) together with (2.18) and (2.16) form a moment system with infinite number of equations, which is equivalent to (2.7).

3 System in 1D Spatial Space

In order to derive the regularization term to achieve the hyperbolicity of the moment systems as in Section 2.3, we first consider in this section the special case with homogeneous dependence of the distribution function on spatial coordinate $\boldsymbol{x}$ except for x_1 direction. Since the velocity space is multi-dimensional, the result in this section is essential different from [1]. The general case in multi-dimensional spatial space is studied in the next section based on the results herein and the Galilean invariance of the regularization.

In 1D spatial space, the distribution function $f(t, x_1, \boldsymbol{\xi})$ satisfies

$$
\frac{\partial f}{\partial t} + \xi_1 \frac{\partial f}{\partial x_1} = 0, \qquad t \in \mathbb{R}^+, \quad x_1 \in \mathbb{R}, \quad \boldsymbol{\xi} \in \mathbb{R}^D.
\tag{3.1}
$$

The moment system in Section 2.3 degenerates to a simpler form. The conservation of mass, momentum and energy (2.8) turn into

$$
\frac{\partial \rho}{\partial t} + u_1 \frac{\partial \rho}{\partial x_1} + \rho\frac{\partial u_1}{\partial x_1} = 0,
\tag{3.2a}
$$

$$
\rho\frac{\partial u_i}{\partial t} + \rho u_1 \frac{\partial u_i}{\partial x_1} + \frac{\partial p_{e_1+e_i}}{\partial x_1} = 0, \quad i = 1, \cdots, D,
\tag{3.2b}
$$

$$
\frac{D}{2}\rho\frac{\partial \theta}{\partial t} + \frac{D}{2}\rho u_1 \frac{\partial \theta}{\partial x_1} + \frac{\partial q_1}{\partial x_1} + \sum_{i=1}^{D} p_{e_1+e_i}\frac{\partial u_i}{\partial x_1} = 0.
\tag{3.2c}
$$

The moment equations (2.16) become

$$
\frac{\partial f_\alpha}{\partial t} + \theta\frac{\partial f_{\alpha-e_1}}{\partial x_1} + u_1 \frac{\partial f_\alpha}{\partial x_1} + (\alpha_1+1)\frac{\partial f_{\alpha+e_1}}{\partial x_1} - \frac{\theta}{2\rho}C_{\theta,\alpha}^{(1)}\frac{\partial \rho}{\partial x_1}
$$
$$
+ \sum_{d=1}^{D}\frac{\partial u_d}{\partial x_1}\left(\theta f_{\alpha-e_d-e_1} + (\alpha_1+1)f_{\alpha-e_d+e_1} - \frac{C_\alpha}{D\rho}p_{e_1+e_d}\right)
$$
$$
+ \sum_{d=1}^{D}\left(-\frac{f_{\alpha-e_d}}{\rho}\frac{\partial p_{e_1+e_d}}{\partial x_1} + \frac{C_{\theta,\alpha}^{(1)}}{2D\rho}\frac{\partial p_{2e_d}}{\partial x_1}\right) - \frac{C_\alpha}{D\rho}\frac{\partial q_1}{\partial x_1} = 0,
\tag{3.3}
$$

where C_α and $C_{\theta,\alpha}^{(1)}$ are defined in (2.17). The governing equations of p_{2e_i} (2.18) turn into: for $i = 1, \cdots, D$,

$$
\frac{\partial p_{2e_i}/2}{\partial t} + u_1 \frac{\partial p_{2e_i}/2}{\partial x_1} + (\frac{1}{2}+\delta_{i1})\rho\theta\frac{\partial u_1}{\partial x_1}
$$
$$
+ \sum_{d=1}^{D}(2\delta_{i1}+1)f_{2e_i-e_d+e_1}\frac{\partial u_d}{\partial x_1} + (2\delta_{i1}+1)\frac{\partial f_{2e_i+e_1}}{\partial x_1} = 0.
\tag{3.4}
$$

Analogously to the moment system in Section 2.1, let $f_\alpha = 0$, $|\alpha| > M$ for $M \geq 3$, and then (3.2) and (3.3), together with (3.4) form a closed moment system corresponding to (3.1).

To facilitate the reading below in studying the moment system, some notations are introduced as follows:

$$\text{if } \boldsymbol{a} = (a_1, \cdots, a_n) \in \mathbb{R}^n, \text{ then } \boldsymbol{a}(i{:}j) = (a_i, \cdots, a_j), \tag{3.5a}$$

$$\text{if } \mathbf{A} = (a_{ij})_{n \times n} \in \mathbb{R}^{n \times n}, \text{ then } \mathbf{A}(i, j{:}k) = (a_{i,j}, \cdots, a_{i,k}), \tag{3.5b}$$

$$\mathbf{A}(i{:}l, \, j{:}k) = \begin{pmatrix} a_{i,j} & a_{i,j+1} & \cdots & a_{i,k} \\ a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{l,j} & a_{l,j+1} & \cdots & a_{l,k} \end{pmatrix}, \tag{3.5c}$$

$$\mathcal{D} = \{1, 2, \cdots, D\}, \qquad\qquad \tilde{\alpha} = (0, \alpha_2, \cdots, \alpha_D), \tag{3.6a}$$

$$\hat{\alpha} = (\alpha_2, \cdots, \alpha_D) \in \mathbb{N}^{D-1}, \qquad\qquad \alpha! = \prod_{i=1}^{D} \alpha_i!, \quad |\alpha| = \sum_{i=1}^{D} \alpha_i, \tag{3.6b}$$

$$\mathcal{S}_{D,M} = \{\alpha \in \mathbb{N}^D \mid |\alpha| \leq M\}, \qquad \mathcal{S}_{D,M}(\hat{\alpha}) = \{\beta \in \mathcal{S}_{D,M} \mid \hat{\beta} = \hat{\alpha}\}. \tag{3.6c}$$

We permute the elements of $\mathcal{S}_{D,M}$ by lexicographic order. Then for any $\alpha \in \mathcal{S}_{D,M}$,

$$\mathcal{N}_D(\alpha) = \sum_{i=1}^{D} \binom{\sum_{k=D-i+1}^{D} \alpha_k + i - 1}{i} + 1 \tag{3.7}$$

holds, where $\mathcal{N}_D(\alpha)$ is the ordinal number of α in $\mathcal{S}_{D,M}$, and the cardinal number of set $\mathcal{S}_{D,M}$ is $N = \mathcal{N}_D(Me_D) = \binom{M+D}{D}$, which is total number of moments if a truncation with $|\alpha| \leq M$ is introduced. In addition, it is clear that for each $\alpha, \beta \in \mathcal{S}_{D,M}$ and $\hat{\alpha} \neq \hat{\beta}$,

$$\mathcal{S}_{D,M}(\hat{\alpha}) \bigcap \mathcal{S}_{D,M}(\hat{\beta}) = \emptyset, \quad \mathcal{S}_{D,M} = \bigcup_{\alpha \in \mathcal{S}_{D,M}} \mathcal{S}_{D,M}(\hat{\alpha}), \tag{3.8}$$

simultaneously hold.

3.1 Structure of coefficient matrix

Similar to the 1D case, a truncation with $|\alpha| \leq M, M \geq 3$ is applied. Let $\boldsymbol{w} \in \mathbb{R}^N$ and for each $i, j \in \mathcal{D}$, and $i \neq j$,

$$w_1 = \rho, \qquad\qquad w_{\mathcal{N}_D(e_i)} = u_i, \tag{3.9a}$$

$$w_{\mathcal{N}_D(2e_i)} = \frac{p_{2e_i}}{2}, \qquad\qquad w_{\mathcal{N}_D(e_i+e_j)} = p_{e_i+e_j}, \tag{3.9b}$$

$$w_{\mathcal{N}_D(\alpha)} = f_\alpha, \quad 3 \leq |\alpha| \leq M. \tag{3.9c}$$

Combining (3.2) with (3.4) and (3.3), we obtain

$$\frac{\partial \boldsymbol{w}}{\partial t} + \mathbf{A}_M \frac{\partial \boldsymbol{w}}{\partial x_1} = 0, \tag{3.10}$$

where $\mathbf{A}_M$ depends on (3.2), (3.4) and (3.3).

Clearly, all the matrix $\mathbf{A}_M$ for any $D \in \mathbb{N}^+$, $3 \leq M \in \mathbb{N}$ are well-defined though quite complex. Here we first give some simple examples and conclude a few basic properties of the matrix $\mathbf{A}_M$.

Example 1. *If $D = 2$, the ordinal number of α in $\mathcal{S}_{D,M}$ is $\mathcal{N}_D(\alpha) = \frac{(\alpha_1+\alpha_2+1)(\alpha_1+\alpha_2)}{2} + \alpha_2 + 1$. The permutation of $\boldsymbol{w}$ is showed in Fig. 2(a). As the simplest case, the matrix $\mathbf{A}_3$ is*

$$
\mathbf{A}_3 = \begin{bmatrix}
u_1 & \rho & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & u_1 & 0 & 2\rho^{-1} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & u_1 & 0 & \rho^{-1} & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{3}{2}p_1 & 0 & u_1 & 0 & 0 & 3 & 0 & 0 & 0 \\
0 & 2f_{1,1} & p_1 & 0 & u_1 & 0 & 0 & 2 & 0 & 0 \\
0 & p_2 & f_{1,1} & 0 & 0 & u_1 & 0 & 0 & 1 & 0 \\
\frac{\rho\theta^2 - 2\theta p_1}{2\rho} & 4f_{3,0} & 0 & \theta & 0 & \frac{2f_{2,0}}{\rho} & u_1 & 0 & 0 & 0 \\
-\frac{3\theta f_{1,1}}{2\rho} & 3f_{2,1} & 3f_{3,0} & -\frac{f_{1,1}}{2\rho} & \frac{\rho\theta - f_{2,0}}{\rho} & \frac{3f_{1,1}}{2\rho} & 0 & u_1 & 0 & 0 \\
-\frac{1}{2}\theta^2 & 2f_{1,2} & 2f_{2,1} & \frac{2f_{2,0}}{\rho} & -\frac{f_{1,1}}{\rho} & \theta & 0 & 0 & u_1 & 0 \\
-\frac{\theta f_{1,1}}{2\rho} & f_{0,3} & f_{1,2} & \frac{f_{1,1}}{2\rho} & \frac{f_{2,0}}{\rho} & \frac{f_{1,1}}{2\rho} & 0 & 0 & 0 & u_1
\end{bmatrix}, \qquad (3.11)
$$

whereas $p_1 = p_{2e_1}$, $f_{m,n} = f_{me_1+ne_2}$. If $M > 3$, for any $\alpha \in \mathbb{N}^2$, and $3 < |\alpha| \le M$,

$$\mathbf{A}_M(1{:}10,\ 1{:}10) = \mathbf{A}_3, \tag{3.12a}$$

$$\mathbf{A}_M(\mathcal{N}_D(\alpha), \mathcal{N}_D(\alpha)) = u_1, \tag{3.12b}$$

$$\mathbf{A}_M(\mathcal{N}_D(\alpha), \mathcal{N}_D(\alpha - e_1)) = \theta, \quad \text{if } \alpha_1 > 0, \tag{3.12c}$$

$$\mathbf{A}_M(\mathcal{N}_D(\alpha), \mathcal{N}_D(\alpha + e_1)) = \alpha_1 + 1, \quad \text{if } |\alpha| < M, \tag{3.12d}$$

$$\mathbf{A}_M(\mathcal{N}_D(\alpha), 1{:}9) = (-\frac{\theta}{2\rho}C_{\theta,\alpha}^{(1)}, \quad \theta f_{\alpha - 2e_1} + (\alpha_1 + 1)f_\alpha - \frac{C_\alpha}{2\rho}p_{2e_1},$$

$$\theta f_{\alpha - e_1 - e_2} + (\alpha_1 + 1)f_{\alpha + e_1 - e_2} - \frac{C_\alpha}{2\rho}p_{e_1 + e_2}, \tag{3.12e}$$

$$-2\frac{f_{\alpha - e_1}}{\rho} + \frac{C_{\theta,\alpha}^{(1)}}{2\rho}, \quad -\frac{f_{\alpha - e_2}}{\rho}, \frac{C_{\theta,\alpha}^{(1)}}{2\rho}, -\frac{3C_\alpha}{2\rho}, \ 0, \quad -\frac{C_\alpha}{2\rho}),$$

where C_α and $C_{\theta,\alpha}^{(1)}$ are defined in equation (2.17). We remark that

- *an entry $\mathbf{A}_M(i,j)$, if not defined above, is taken as zero;*

- *for $|\alpha| = 4$, some $\mathbf{A}_M(i,j)$ may be double defined in (3.12c) and (3.12e), the value of which is the sum of the both expressions.*

Fig. 1 gives the sparse matrix pattern of $\mathbf{A}_M$ with $M = 8$ and $D = 2$. It is observed that there is no more than one nonzero component of $\mathbf{A}_M(i, i + 1 : N)$, for each $i = 1$, $\cdots$, N. Precisely, there is a unique nonzero component as $1 \le i \le \mathcal{N}_D((M - 1)e_D)$ and $\mathbf{A}_M(i, i + 1 : N) = \mathbf{0}$ as $i > \mathcal{N}_D((M - 1)e_D)$. Noticing that the column index of the nonzero entries in the upper triangular part of $\mathbf{A}_M$ on different rows are different from each other, this makes one recall the form of lower Hessenberg matrix, of which the only nonzero entries in the upper triangular part on the i-th row is located at position $(i, i+1)$. This property of Hessenberg matrix makes it very convenient for one to calculate
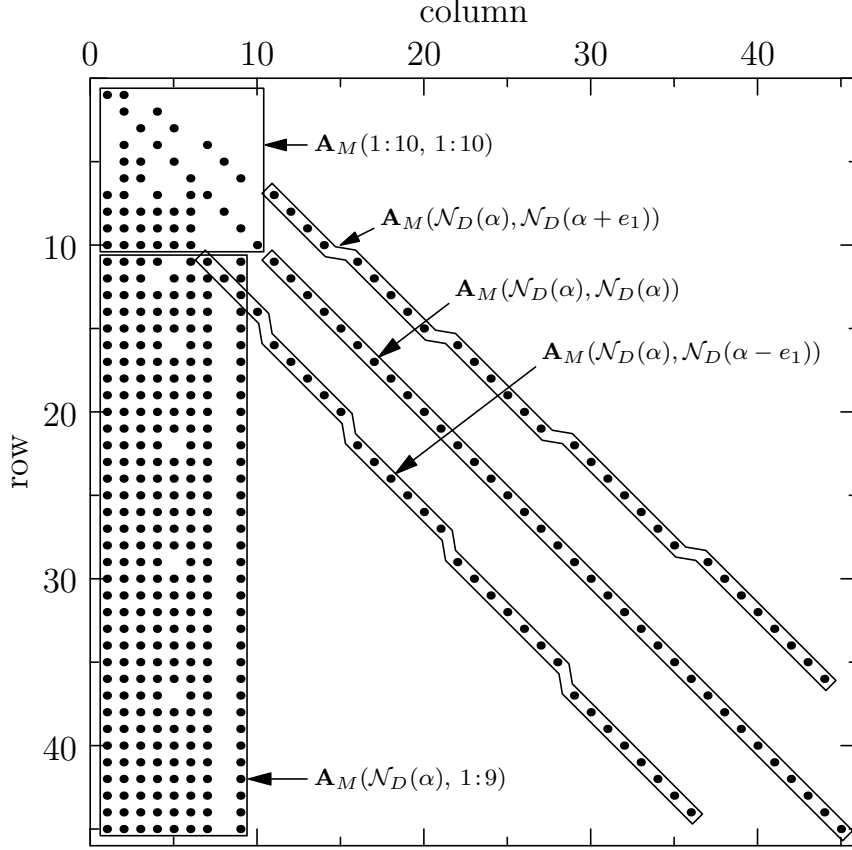
Figure 1: The sparse matrix pattern of $\mathbf{A}_M$ with $M = 8, D = 2$. Its nonzero entries are defined as in (3.12).

its eigenvectors once eigenvalues are given using a row by row sequential procedure. Here $\mathbf{A}_M$ is essentially the same as lower Hessenberg matrix on this point, and we notice that its lower triangular part is sparse, hence we are provided the approach to calculate the eigenvalues $\mathbf{A}_M$ together with the corresponding eigenvectors using the same technique.

Furthermore, (3.12) shows the diagonal entries of the matrix $\mathbf{A}_M$ are all u_1, and the entries of the matrix $\mathbf{A}_M - u_1 \mathbf{I}$ are independent of $u_i, i \in \mathcal{D}$, where $\mathbf{I}$ is the $N \times N$ identity matrix. In fact, (3.10) can be written as

$$\frac{\mathrm{D}\boldsymbol{w}}{\mathrm{D}t} + (\mathbf{A}_M - u_1 \mathbf{I}) \frac{\partial \boldsymbol{w}}{\partial x_1} = 0, \tag{3.13}$$

where $\dfrac{\mathrm{D}}{\mathrm{D}t}$ is material derivative defined as

$$\frac{\mathrm{D}}{\mathrm{D}t} = \frac{\partial}{\partial t} + u_1 \frac{\partial}{\partial x_1}.$$

Hence, that $\mathbf{A}_M - u_1 \mathbf{I}$ are independent of $u_i, i \in \mathcal{D}$ indicates the moment system is translation invariant. On the other hand, eigenvalues of $\mathbf{A}_M$ can be written in the form $u_1 + a$, where a is indeterminate and independent of $u_i, \ i \in \mathcal{D}$, and eigenvectors of it is independent of $u_i, \ i \in \mathcal{D}$, too.

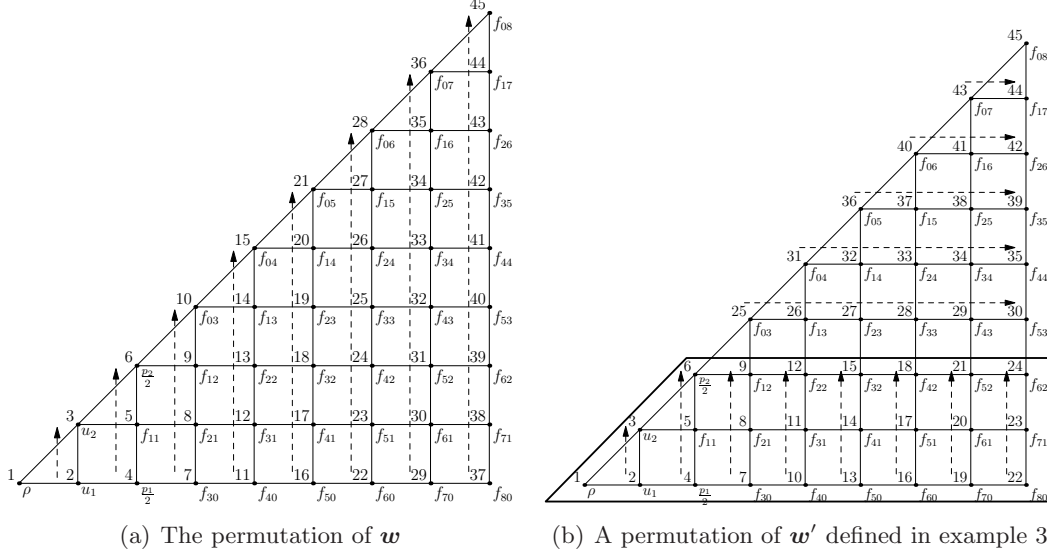(a) The permutation of $\boldsymbol{w}$ (b) A permutation of $\boldsymbol{w}'$ defined in example 3

Figure 2: The permutation of moments while $D = 2, M = 8$. Each node stands for one moment. The marks in the lower right of the node shows the expression of the moment, while the number in the upper left represents the ordinal number in $\boldsymbol{w}$ or $\boldsymbol{w}'$. The dashed arrows depict the path of the corresponding permutation. The left one is the permutation of $\boldsymbol{w}$, and the right one is a permutation of $\boldsymbol{w}'$ defined in example 3.

Example 2. *Considering the case $D = 2$, we can write out the matrix $\mathbf{A}_M$ according to example 1, for any $3 \leq M \in \mathbb{N}$. If we let $f_\alpha = 0$ for all $\alpha \in \mathbb{N}^2$, and $|\alpha| \leq M$ except f_0 and f_{Me_1}, direct calculation gives the characteristic polynomial of $\mathbf{A}_M$ as*

$$
\begin{aligned}
|\lambda \mathbf{I} - \mathbf{A}_M| = {} & \left(\prod_{i=1}^{M-1} He_i \left(\frac{\lambda - u_1}{\sqrt{\theta}} \right) \theta^{i/2} \right) \\
& \times \left(He_M \left(\frac{\lambda - u_1}{\sqrt{\theta}} \right) \theta^{M/2} + (-1)^{M-1} M! f_{Me_1} \right) \\
& \times \left(He_{M+1} \left(\frac{\lambda - u_1}{\sqrt{\theta}} \right) \theta^{(M+1)/2} + (-1)^{M-1} (M+1)! f_{Me_1} (\lambda - u_1) \right).
\end{aligned}
\tag{3.14}
$$

The matrix $\mathbf{A}_M$, obviously, has complex eigenvalues, for some f_{Me_1}.

Analogously, involved calculations with the help of computer algebraic system show that the matrix $\mathbf{A}_M$ has complex eigenvalues for some admissible $\boldsymbol{w}$, for any $D \geq 3$. This reveals that $\mathbf{A}_M$ is not diagonalisable with real eigenvalues for some $\boldsymbol{w}$ as $D = 1$ in [1].

Remark 1. Since moments f_α are related to $f(t, \boldsymbol{x}, \boldsymbol{\xi})$ by (2.2), the moments f_α can not be arbitrary number if the distribution function is to be kept positive. Particularly, ρ and θ given by (2.5) clearly satisfy

$$
\rho > 0, \quad \theta > 0.
\tag{3.15}
$$

Though (3.15) is not enough to provide us a positive $f(t, \boldsymbol{x}, \boldsymbol{\xi})$, the discussion in this paper requires no further constraints on the other moments. Hence, in this paper the admissible $\boldsymbol{w}$ stands for $\boldsymbol{w}$ which satisfies (3.15).

Example 3. *Actually, it can be observed that the matrix $\mathbf{A}_M$ is reducible if we rearrange $\boldsymbol{w}$ as $\boldsymbol{w}'$ using another permutation rule. In case of $D = 2$, the rule reads:*

10

1. The moments with $\alpha_2 \leq 2$ are arranged at first using the lexicographic order;

2. The rest moments are arranged then using the lexicographic order based on index transformed as (α_2, α_1).

Clearly, $\boldsymbol{w}$ and $\boldsymbol{w}'$ are related by a permutation matrix $\mathbf{P}$ that $\boldsymbol{w}' = \mathbf{P}\boldsymbol{w}$. The Fig. 2(b) gives a schematic diagram of the permutation rule for $\boldsymbol{w}'$ with $M = 8$. Let $\mathbf{A}'_M = \mathbf{P}\mathbf{A}_M\mathbf{P}^{-1}$, then

$$\frac{\partial \boldsymbol{w}'}{\partial t} + \mathbf{A}'_M \frac{\partial \boldsymbol{w}'}{\partial x_1} = 0, \tag{3.16}$$

holds. Fig. 3.1 gives the sparse matrix pattern of $\mathbf{A}'_M$. It is obvious that $\mathbf{A}'_M$ is reducible (see e.g. [6] for the definition), and can be reduced into $M-1$ blocks. $\mathcal{S}_{D,M}(\hat{e}_1) \bigcup \mathcal{S}_{D,M}(\hat{e}_2)$ $\bigcup \mathcal{S}_{D,M}(2\hat{e}_2)$ is one of the blocks, and $\mathcal{S}^D_M(\hat{\alpha})$, for each $\hat{\alpha} \in \mathbb{N}$ and $\hat{\alpha} \neq \hat{e}_1, \hat{e}_2, 2\hat{e}_2$ is another block.
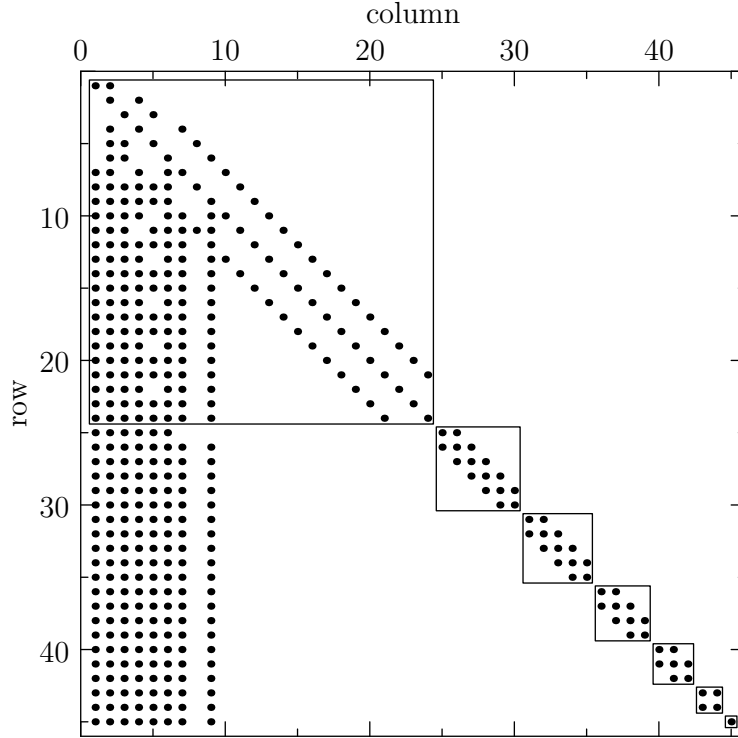

Figure 3: The sparse matrix pattern of $\mathbf{A}'_M$ with $M = 8, D = 2$. $\mathbf{A}'_M$ is reducible.

These examples show some very useful properties of the matrix $\mathbf{A}_M$ as follows:

Property 1. *Matrix $\mathbf{A}_M$ satisfies the following properties:*

1(1). For each $\alpha \in \mathbb{N}^D$, $|\alpha| \leq M$, let $i = \mathcal{N}_D(\alpha)$, then there are no more than one entries of $\mathbf{A}_M(i, i+1{:}N)$ nonzero. In fact, there is a unique nonzero component as $|\alpha| < M$ and $\mathbf{A}_M(i, i+1{:}N) = 0$ as $|\alpha| = M$.

1(2). The diagonal entries of the matrix $\mathbf{A}_M$ are all u_1, and entries of the matrix $\mathbf{A}_M - u_1\mathbf{I}$ are independent of $u_i, i \in \mathcal{D}$.

1(3). $\mathbf{A}_M(\boldsymbol{w})$ may be not diagonalisable with real eigenvalues for some admissible $\boldsymbol{w}$.

11

1(4). $\mathbf{A}_M$ *is reducible, and can be reduced into* $\binom{M+D-1}{D-1} - 2(D-1)$ *blocks.* $\mathcal{S}_{D,M}(\hat{e}_1) \bigcup$ $\left(\bigcup_{k=2}^{D} \mathcal{S}_{D,M}(\hat{e}_k)\right) \bigcup \left(\bigcup_{k=2}^{D} \mathcal{S}_{D,M}(2\hat{e}_k)\right)$ *is one of the blocks, and* $\mathcal{S}_{D,M}(\hat{\alpha})$*, for each* $\hat{\alpha} \in \mathbb{N}^{D-1}$ *and* $\hat{\alpha} \neq \hat{e}_1, \hat{e}_k, 2\hat{e}_k, k = 2, \cdots, D$ *is one of the blocks.*

3.2 Globally hyperbolic regularization

For 1D case, the regularization as (2.13) was proposed such that the moment system turns out to be globally hyperbolic (see [1] for details). Actually, the regularization therein can be extend to multiple dimensional systems. For $D \in \mathbb{N}^+$, let us start from the definition as below:

Definition 1. *For any* $|\alpha| = M$*, let*

$$\mathcal{R}_{M,D}(\alpha) \triangleq \sum_{j=1}^{D} \mathcal{R}_{M,D}^{j}(\alpha), \tag{3.17}$$

where

$$\mathcal{R}_{M,D}^{j}(\alpha) = \sum_{d=1}^{D} f_{\alpha - e_d + e_j} \frac{\partial u_d}{\partial x_j} + \frac{1}{2}\left(\sum_{d=1}^{D} f_{\alpha - 2e_d + e_j}\right)\frac{\partial \theta}{\partial x_j}. \tag{3.18}$$

As in 1D case [1], $\mathcal{R}_{M,D}(\alpha)$ *is the regularization terms based on the characteristic speed correction.*

With (2.15), we have that

$$\mathcal{R}_{M,D}^{j}(\alpha) = \sum_{d=1}^{D} f_{\alpha - e_d + e_j} \frac{\partial u_d}{\partial x_j} + \left(\sum_{d=1}^{D} f_{\alpha - 2e_d + e_j}\right)\left(\sum_{i=1}^{D} \frac{1}{D\rho}\frac{\partial p_{2e_i}/2}{\partial x_j} - \frac{\theta}{2\rho}\frac{\partial \rho}{\partial x_j}\right).$$

For the case that the dependence of f on $\boldsymbol{x}$ is only on x_1, we have that

$$\mathcal{R}_{M,D}^{j}(\alpha) = 0, \quad \text{for} \ \ j = 2, \cdots, D.$$

This leads to $\mathcal{R}_{M,D}(\alpha) = \mathcal{R}_{M,D}^{1}(\alpha)$. The regularized system is obtained by subtracting $\mathcal{R}_{M,D}(\alpha)$ from the governing equation of f_α in (3.10), for $|\alpha| = M$.

Definition 2. $\hat{\mathbf{A}}_M$ *is called the regularized matrix of the matrix* $\mathbf{A}_M$*, if it satisfies that for any admissible* $\boldsymbol{w}$*,*

$$\hat{\mathbf{A}}_M \frac{\partial \boldsymbol{w}}{\partial x_1} = \mathbf{A}_M \frac{\partial \boldsymbol{w}}{\partial x_1} - \sum_{|\alpha|=M} \mathcal{R}_{M,D}^{1}(\alpha) I_{\mathcal{N}_D(\alpha)}, \tag{3.19}$$

where I_k *is the* k*-th column of the* $N \times N$ *identity matrix.*

The regularization terms only change a few entries of the lower triangular part of $\mathbf{A}_M$, with the order of the corresponding moments equals to M, so that the Properties 1(1), 1(2), 1(4) of $\mathbf{A}_M$ are also valid for $\hat{\mathbf{A}}_M$, while the Property 1(3) is changed to be the diagonalisability of $\hat{\mathbf{A}}_M$ over the real field $\mathbb{R}$. Actually, we have the following theorem:

Theorem 1. *The regularized moment system*

$$\frac{\partial \boldsymbol{w}}{\partial t} + \hat{\mathbf{A}}_M \frac{\partial \boldsymbol{w}}{\partial x_1} = 0 \tag{3.20}$$

is hyperbolic for any admissible $\boldsymbol{w}$.

The definition of the hyperbolicity shows that this theorem is equivalent to the diagonalisability of $\hat{\mathbf{A}}_M$ with real eigenvalues for any admissible $\boldsymbol{w}$. Before proving this result, we first make some simplifications and give several useful lemmas.

Let us denote that

$$\boldsymbol{d} = (d_j)_{N \times 1}, \qquad\qquad d_1 = \rho^{-1}, \tag{3.21a}$$

$$d_{j+1} = \theta^{-1/2}, \quad j = \mathcal{D}, \qquad d_{\mathcal{N}_D(\alpha)} = \rho^{-1}\theta^{-|\alpha|/2}, \quad 2 \le |\alpha| \le M, \tag{3.21b}$$

$$\boldsymbol{\Lambda} = \operatorname{diag}\{d_1, d_2, \cdots, d_N\}, \tag{3.21c}$$

$$\hat{p}_{e_i+e_k} = \frac{p_{e_i+e_k}}{\rho\theta}, \quad i,k = \mathcal{D}, \qquad g_\alpha = \frac{f_\alpha}{\rho\theta^{|\alpha|}}. \tag{3.21d}$$

By the virtue of Property 1(2), we let

$$\hat{\mathbf{A}}_M = u_1 \mathbf{I} + \sqrt{\theta}\,\boldsymbol{\Lambda}^{-1}\tilde{\mathbf{A}}_M\boldsymbol{\Lambda}. \tag{3.22}$$

Then, we can obtain properties of $\hat{\mathbf{A}}_M$ by studying the matrix $\tilde{\mathbf{A}}_M$. Since $\tilde{\mathbf{A}}_M$ is related to $\hat{\mathbf{A}}_M - u_1\mathbf{I}$ by a similarity transformation by a diagonal matrix, Property 1(1) holds for $\tilde{\mathbf{A}}_M$. Hence, it is convenient to calculate the eigenvectors of matrix $\tilde{\mathbf{A}}_M$. Firstly, we denote by some symbols as

$$r_{w_{\mathcal{N}_D(\alpha)}} = v_{\mathcal{N}_{D-1}(\hat{\alpha})}, \qquad\qquad \text{if } \alpha_1 = 0, \tag{3.23}$$

$$r_{u_1} = \lambda r_\rho, \qquad\qquad r_{p_{2e_1}/2} = \frac{\lambda^2}{2}r_\rho, \tag{3.24}$$

$$r_{f_{e_i}} = r_{f_\alpha} = 0, \qquad\qquad \text{if at least one } \alpha_j < 0, j \in \mathcal{D} \tag{3.25}$$

$$r_{p_{e_1+e_k}} = \lambda r_{u_k}, \qquad\qquad k \in \mathcal{D}\backslash\{1\}, \tag{3.26}$$

$$r_{f_{2e_i}} = r_{p_{2e_i}/2} - \sum_{d=1}^{D} \frac{r_{p_{2e_d}/2}}{D}, \tag{3.27}$$

$$r_{w_{\mathcal{N}_D(\alpha)}} = \frac{He_{\alpha_1}(\lambda)}{\alpha_1!}\big(r_{f_{\tilde{\alpha}}} + G(\tilde{\alpha})\big) - G(\alpha), \quad \text{if } \alpha_1 \ne 0,\ |\alpha| \ge 3. \tag{3.28}$$

where $v_{\mathcal{N}_{D-1}(\hat{\alpha})}$ and λ are indeterminate parameters, and

$$G(\alpha) = \sum_{d=1}^{D} g_{\alpha-e_d}r_{u_d} + \left(\sum_{d=1}^{D} \frac{r_{p_{2e_d}/2}}{D} - \frac{r_\rho}{2}\right)\sum_{k=1}^{D} g_{\alpha-2e_k}. \tag{3.29}$$

For better readability, here we adopt the notations [e.g. w_k, f_{e_i}, f_{2e_i}] as the subscript of r. This does not mean the subscripts are taken as the value of them, but only taken literally as the notations themselves. We collect these r_{w_k}, with w_k as a component of $\boldsymbol{w}$ and $k = 1, \cdots, N$, to produce a vector $\boldsymbol{r} \in \mathbb{R}^N$ as

$$\boldsymbol{r} = (r_{w_1}, r_{w_2}, \cdots, r_{w_N}), \tag{3.30}$$

where N is total number of moments. Here it is clear that $\boldsymbol{r}$ is prescribed once $v_{\mathcal{N}_{D-1}(\hat{\alpha})}$ and λ are all given. With particular setup of these parameters, λ and $\boldsymbol{r}$ is turned out to be a pair of eigenvalue and eigenvector of $\tilde{\mathbf{A}}_M$. Precisely, we have the following lemma:

13

Lemma 1. $r \neq 0$ *is the right eigenvector of the matrix* $\tilde{\mathbf{A}}_M$ *for the eigenvalue* λ *if*

$$\frac{He_{\alpha_1+1}(\lambda)}{(\alpha_1+1)!}\big(r_{w_{\tilde{\alpha}}} + G(\tilde{\alpha})\big) = 0 \tag{3.31}$$

holds, for all $|\alpha| = M$.

Proof. Let $i = \mathcal{N}_D(\alpha)$, with $|\alpha| \leq M$, then we need only to verify

$$\tilde{\mathbf{A}}_M(i, 1:N) \cdot \boldsymbol{r} = \lambda r_{w_i} \tag{3.32}$$

always valid. Since $\mathbf{A}_M$ is determined by (3.2), (3.3) and (3.4), and $\hat{\mathbf{A}}_M$ and $\tilde{\mathbf{A}}_M$ are defined as (3.19) and (3.22), respectively, we can write any entries of $\tilde{\mathbf{A}}_M$. Now let us verify the equation (3.32) case by case:

- For $\alpha = 0$,
$$\tilde{\mathbf{A}}_M(i, 1:N) \cdot \boldsymbol{r} = 1 \cdot r_{u_1} = \lambda r_\rho. \tag{3.33}$$

- For $\alpha = e_1$,
$$\tilde{\mathbf{A}}_M(i, 1:N) \cdot \boldsymbol{r} = 2 \cdot r_{p_{2e_1}/2} = \lambda^2 r_\rho = \lambda r_{u_1}. \tag{3.34}$$

- For $\alpha = e_k$, $k = 2, \cdots, D$,
$$\tilde{\mathbf{A}}_M(i, 1:N) \cdot \boldsymbol{r} = 1 \cdot r_{p_{e_1+e_k}/2} = \lambda r_{u_k}. \tag{3.35}$$

- For $\alpha = 2e_1$,

$$\begin{aligned}
\tilde{\mathbf{A}}_M(i, 1:N) \cdot \boldsymbol{r} &= \frac{3}{2}r_{u_1} + \sum_{d=1}^{D} 3g_{\alpha+e_1-e_d}r_{u_d} + 3r_{f_{3e_1}} \\
&= \frac{3}{2}r_{u_1} + 3g_{2e_1}r_{u_1} + 3\left(\frac{He_3(\lambda)}{6}r_\rho - g_{2e_1}r_{u_1}\right) \\
&= \frac{\lambda^3}{2}r_\rho = \lambda r_{p_{2e_1}/2}.
\end{aligned} \tag{3.36}$$

- For $\alpha = 2e_k$, $k = 2, \cdots, D$

$$\begin{aligned}
&\tilde{\mathbf{A}}_M(i, 1:N) \cdot \boldsymbol{r} \\
&= \frac{1}{2}r_{u_1} + \sum_{d=1}^{D} g_{\alpha+e_1-e_d}r_{u_d} + r_{f_{e_1+2e_k}} \\
&= \frac{1}{2}r_{u_1} + g_\alpha r_{u_1} + g_{e_1+e_k}r_{u_k} + \left(\lambda\big(r_{p_{2e_k}/2} - \frac{1}{2}r_\rho\big) - g_\alpha r_{u_1} - g_{e_1+e_k}r_{u_k}\right) \\
&= \lambda r_{p_{2e_k}/2}.
\end{aligned} \tag{3.37}$$

- For $\alpha = e_1 + e_k$, $k = 2, \cdots, D$

$$\begin{aligned}
&\tilde{\mathbf{A}}_M(i, 1:N) \cdot \boldsymbol{r} \\
&= \frac{1}{2}r_{u_1} + \sum_{d=1}^{D} g_{\alpha+e_1-e_d}r_{u_d} + r_{f_{e_1+2e_k}} \\
&= \frac{1}{2}r_{u_1} + g_\alpha r_{u_1} + g_{e_1+e_k}r_{u_k} + \left(\lambda\big(r_{p_{2e_k}/2} - \frac{1}{2}r_\rho\big) - g_\alpha r_{u_1} - g_{e_1+e_k}r_{u_k}\right) \\
&= \lambda r_{p_{2e_k}/2}.
\end{aligned} \tag{3.38}$$

- For $3 \leq |\alpha| < M$, $\alpha_1 > 0$

$$
\begin{aligned}
\tilde{\mathbf{A}}_M(i, 1:N) \cdot \boldsymbol{r} &= 1 \cdot r_{f_{\alpha-e_1}} + (\alpha_1 + 1) r_{f_{\alpha+e_1}} - \frac{1}{2} \tilde{C}^{(1)}_{\theta,\alpha} r_\rho \\
&+ \sum_{d=1}^{D} \left(g_{\alpha-e_d-e_1} + (\alpha_1 + 1) g_{\alpha-e_d+e_1} - \frac{\tilde{C}_\alpha}{D} \hat{p}_{e_1+e_d} \right) r_{u_d} \\
&+ \sum_{d=1}^{D} \left(-g_{\alpha-e_d} r_{p_{e_1+e_d}} + \frac{\tilde{C}^{(1)}_{\theta,\alpha}}{D} r_{p_{2e_d}/2} \right) - \frac{\tilde{C}_\alpha}{D} r_{q_1} \\
&\triangleq X_1 + X_2,
\end{aligned}
\tag{3.39}
$$

whereas

$$
\tilde{C}_\alpha = \sum_{k=1}^{D} g_{\alpha-2e_k},
$$

$$
\tilde{C}^{(1)}_{\theta,\alpha} = \sum_{k=1}^{D} \left(g_{\alpha-2e_k-e_1} + (\alpha_1 + 1) g_{\alpha-2e_k+e_1} \right),
$$

$$
r_{q_1} = 3 r_{f_{3e_1}} + \sum_{k=2}^{D} r_{f_{e_1+2e_k}} = -\frac{D}{2} \lambda r_\rho + \sum_{k=1}^{D} \left(\lambda r_{p_{2e_k}/2} - \hat{p}_{e_1+e_k} r_{u_k} \right),
$$

$$
X_1 = r_{f_{\alpha-e_1}} + (\alpha_1 + 1) r_{f_{\alpha+e_1}}, \quad X_2 \text{ is the rest terms.}
$$

Substituting (3.29) into X_1 yields

$$
\begin{aligned}
X_1 &= \frac{He_{\alpha_1-1}(\lambda)}{(\alpha_1-1)!} \left(r_{f_{\tilde{\alpha}}} + G(\tilde{\alpha}) \right) - G(\alpha - e_1) \\
&+ (\alpha_1 + 1) \left(\frac{He_{\alpha_1+1}(\lambda)}{(\alpha_1+1)!} \left(r_{f_{\tilde{\alpha}}} + G(\tilde{\alpha}) \right) - G(\alpha + e_1) \right) \\
&= \lambda \frac{He_{\alpha_1}(\lambda)}{\alpha_1!} \left(r_{f_{\tilde{\alpha}}} + G(\tilde{\alpha}) \right) - G(\alpha - e_1) - (\alpha_1 + 1) G(\alpha + e_1).
\end{aligned}
\tag{3.40}
$$

For X_2, using (3.26), we get

$$
\begin{aligned}
X_2 &= -\frac{r_\rho}{2} \sum_{k=1}^{D} \left(g_{\alpha-e_1-2e_k} + (\alpha_1 + 1) g_{\alpha+e_1-2e_k} - \lambda g_{\alpha-2e_k} \right) \\
&+ \sum_{d=1}^{D} r_{u_d} \left(g_{\alpha-e_1-e_d} + (\alpha_1 + 1) g_{\alpha+e_1-e_d} - \lambda g_{\alpha-e_d} \right) \\
&+ \left(\sum_{d=1}^{D} \frac{r_{p_{2e_d}/2}}{D} \right) \sum_{k=1}^{D} \left(g_{\alpha-e_1-2e_k} + (\alpha_1 + 1) g_{\alpha+e_1-2e_k} - \lambda g_{\alpha-2e_k} \right).
\end{aligned}
\tag{3.41}
$$

15

Now we calculate $G(\alpha - e_1) + (\alpha_1 + 1)G(\alpha + e_1) - \lambda G(\alpha)$. Some simplification gives

$$
\begin{aligned}
& G(\alpha - e_1) + (\alpha_1 + 1)G(\alpha + e_1) - \lambda G(\alpha) \\
&= \sum_{d=1}^{D} r_{u_d} \left(g_{\alpha - e_1 - e_d} + (\alpha_1 + 1)g_{\alpha + e_1 - e_d} - \lambda g_{\alpha - e_d} \right) \\
&\quad + \left(\sum_{d=1}^{D} \frac{r_{p2e_d}/2}{D} - \frac{r_\rho}{2} \right) \sum_{k=1}^{D} \left(g_{\alpha - e_1 - 2e_k} + (\alpha_1 + 1)g_{\alpha + e_1 - 2e_k} - \lambda g_{\alpha - 2e_k} \right) \\
&= X_2.
\end{aligned}
\tag{3.42}
$$

(3.40), (3.41) and (3.42) show

$$
X_1 + X_2 = \lambda \frac{He_{\alpha_1}(\lambda)}{\alpha_1!} \left(r_{f_{\tilde\alpha}} + G(\tilde\alpha) \right) - \lambda G(\alpha) = \lambda r_{w_i}.
\tag{3.43}
$$

- For $3 \le |\alpha| < M$, $\alpha_1 = 0$ or $\alpha = e_k + e_j$, $j > k > 1$, it is the case that to let $r_{f_{\alpha - e_1}} = 0$ in (3.39), which is actually part of (3.25). Hence, (3.32) is valid in this case.

- For $|\alpha| = M$, if $\alpha_1 > 0$, then (3.19) and (3.22) show this case equals to let

$$
r_{f_{\alpha + e_1}} + \sum_{d=1}^{D} g_{\alpha - e_d + e_1} r_{u_d} + \left(\sum_{d=1}^{D} g_{\alpha - 2e_d + e_1} \right) \left(\sum_{i=1}^{D} \frac{1}{D} r_{p2e_i}/2 - \frac{1}{2} r_\rho \right) = 0
\tag{3.44}
$$

in (3.39). Since

$$
G(\alpha + e_1) = \sum_{d=1}^{D} g_{\alpha - e_d + e_1} r_{u_d} + \left(\sum_{d=1}^{D} g_{\alpha - 2e_d + e_1} \right) \left(\sum_{i=1}^{D} \frac{1}{D} r_{p2e_i}/2 - \frac{1}{2} r_\rho \right),
$$

we need only to prove

$$
r_{f_{\alpha + e_1}} + G(\alpha + e_1) = 0.
$$

Actually, it is what (3.31) tells.

If $\alpha_1 = 0$, (3.19) and (3.22) show this case equals to let $r_{f_{\alpha - e_1}} = 0$ and (3.44) valid in (3.39). The former is part of (3.25), while the latter is proved above. Hence, (3.32) is valid in this case.

Collecting all the cases above, we conclude (3.32) is valid for arbitrary α. The lemma is proved. $\qquad\square$

For any α, let $\beta = \alpha + ke_1, k \in \mathbb{N}$, then $\tilde\alpha = \tilde\beta$ holds. Therefore, $r_{w_{\tilde\alpha}}, |\alpha| \le M$ is equivalent to $r_{w_{\tilde\alpha}}, |\alpha| = M$. Hence, in the lemma 1, parameters $r_{\tilde\alpha}, |\alpha| = M$ and λ, are all indeterminate. Let $N_v = \mathcal{N}_{D-1}(M\hat e_D)$, and

$$
\boldsymbol{v} = (v_1, v_2, \cdots, v_{N_v}) \in \mathbb{R}^{N_v}.
$$

Since $\boldsymbol{r}$ is determined by $\boldsymbol{v}$ and λ, by studying the space of the parameters $\boldsymbol{v}$ and λ, we can fully clarify the structure of the eigenvectors of $\tilde{\mathbf{A}}_M$. We have the following lemma that

16

Lemma 2. $\tilde{\mathbf{A}}_M$ *has N linearly independent eigenvectors.*

Proof. For $|\alpha| = M$, (3.31) can be written as

- if $\tilde{\alpha} = 0$, then $\mathcal{N}_{D-1}(\hat{\alpha}) = 1$ and

$$v_1 He_{M+1}(\lambda) = 0, \tag{3.45}$$

- if $\tilde{\alpha} = e_k$, $k \in \mathcal{D}\backslash\{1\}$, then $\mathcal{N}_{D-1}(\hat{\alpha}) = k$ and

$$v_k He_M(\lambda) = 0, \tag{3.46}$$

- if $\tilde{\alpha} = 2e_k$, $k \in \mathcal{D}\backslash\{1\}$, then

$$
\begin{aligned}
0 &= \left(v_{\mathcal{N}_{D-1}(2\hat{e}_k)} + \sum_{d=1}^{D} \frac{r_{w_{\mathcal{N}_D(2e_d)}}}{D} - \frac{v_1}{2} \right) He_{M-1}(\lambda) \\
&= \left(v_{\mathcal{N}_{D-1}(2\hat{e}_k)} + \sum_{d=2}^{D} \frac{v_{\mathcal{N}_{D-1}(2\hat{e}_d)}}{D} + \frac{\lambda^2}{2D}v_1 - \frac{v_1}{2} \right) He_{M-1}(\lambda),
\end{aligned}
\tag{3.47}
$$

- if $\tilde{\alpha} = e_k + e_l$, $k \neq l$, and $k, l \in \mathcal{D}\backslash\{1\}$, then

$$v_{\mathcal{N}_{D-1}(\hat{\alpha})} He_{M-1}(\lambda) = 0, \tag{3.48}$$

- otherwise $(3 \leq |\tilde{\alpha}| \leq M)$,

$$\left(v_{\mathcal{N}_{D-1}(\hat{\alpha})} + \sum_{d=2}^{D} g_{\tilde{\alpha}-e_d} v_d + \sum_{i=1}^{D} g_{\tilde{\alpha}-2e_i} \left(\sum_{d=2}^{D} \frac{v_{\mathcal{N}_{D-1}(2e_d)}}{D} + \frac{\lambda^2}{2D}v_1 - \frac{v_1}{2} \right) \right) He_{\alpha_1+1}(\lambda) = 0. \tag{3.49}$$

Let

$$\mathbf{z}_\lambda = (He_{M+1}(\lambda), \ \cdots, \ \underbrace{He_k(\lambda), \ \cdots, \ He_k(\lambda)}_{\binom{D-1+M-k}{M+1-k} \text{ entries}}, \ \cdots, \ He_1(\lambda)), \tag{3.50}$$

where the $\mathcal{N}_{D-1}(\hat{\alpha})$-th component of $\mathbf{z}_\lambda$ is $He_{\alpha_1+1}(\lambda)$, and the cardinal number of set

$$\#\{\alpha \mid |\alpha| = M, \ \alpha_1 = k - 1\} = \binom{D-1+M-k}{M+1-k}.$$

Equations (3.45), (3.46), (3.47), (3.48) and (3.49) can be collected as

$$\mathbf{z}_\lambda \circ \mathbf{B}\mathbf{v} = 0, \tag{3.51}$$

where $\mathbf{c} = \mathbf{a} \circ \mathbf{b}$, stands for $c_i = a_i b_i$, $i = 1, \cdots, n$, and $\mathbf{B}$ is a $(N_v + 1) \times (N_v + 1)$ real matrix. Precisely, the formation of $\mathbf{B}$ is as

$$
\mathbf{B} = \begin{bmatrix} \mathbf{I} & 0 & 0 \\ \mathbf{B}_{21} & \mathbf{B}_{22} & 0 \\ \mathbf{B}_{31} & \mathbf{B}_{32} & \mathbf{I} \end{bmatrix}
\begin{array}{l} \leftarrow \quad D \text{ row} \\ \leftarrow \quad D(D-1)/2 \text{ row} \\ \leftarrow \quad N_v - D(D+1)/2 \text{ row} \end{array}
\tag{3.52}
$$

$$
\begin{array}{ccc}
\uparrow & \uparrow & \uparrow \\
D \text{ col} & \frac{D(D-1)}{2} \text{ col} & N_v - \frac{D(D+1)}{2} \text{ col}
\end{array}
$$

where $\mathbf{I}$ is identity matrix, whose dimension is context depended. The first D rows of $\mathbf{B}$ are arising from (3.45) and (3.46), the following $D(D-1)/2$ rows are arising from (3.47) and (3.48), and the rest $N_v - D(D+1)/2$ rows are arising from (3.49).

The properties of $\mathbf{B}$ are here further clarified. We denote the entry of $\mathbf{B}$ located at (i,j) position as b_{ij}, whereas $i,j = 0, \cdots, N_v$. Noticing that entries of $\mathbf{B}_{21}$ are from (3.47) and (3.48), we have

$$
b_{ij} = \begin{cases} \dfrac{\lambda^2}{2D} - \dfrac{1}{2}, & \text{if } j = 1, \text{ and } i = \mathcal{N}_{D-1}(2\hat{e}_l) \text{ for some } l = 2, \cdots, D, \\ 0, & \text{else.} \end{cases} \tag{3.53}
$$

And (3.47) and (3.48) make that $\mathbf{B}_{22}$ is a summation as

$$
\mathbf{B}_{22} = \mathbf{I} + \frac{1}{D}\boldsymbol{\Omega}, \tag{3.54}
$$

where $\boldsymbol{\Omega} \in \mathbb{R}^{\frac{D(D-1)}{2} \times \frac{D(D-1)}{2}}$ is a matrix with its entries ω_{ij} as

$$
\omega_{ij} = \begin{cases} 1, & \text{if } i = \mathcal{N}_{D-1}(2\hat{e}_k) - D, \ j = \mathcal{N}_{D-1}(2\hat{e}_l) - D \text{ for some } k,l \in \mathcal{D}\backslash\{1\}, \\ 0, & \text{else.} \end{cases}
$$

Since there are at most $D - 1$ nonzero entries in each row of $\boldsymbol{\Omega}$, it is clear that $\mathbf{B}_{22}$ is strictly diagonally dominant thus nonsingular, and then $\mathbf{B}$ is nonsingular. With the entry value of $\boldsymbol{\Omega}$ as given above, one can check $\boldsymbol{\Omega}^2 = (D-1)\boldsymbol{\Omega}$ holds. Hence we can get the inverse of $\mathbf{B}_{22}$, which reads:

$$
\mathbf{B}_{22}^{-1} = \mathbf{I} - \frac{1}{2D-1}\boldsymbol{\Omega}. \tag{3.55}
$$

Meanwhile, since $\mathbf{B}$ is a nonsingular block lower triangular matrix, we can get its inverse as

$$
\mathbf{B}^{-1} = \begin{bmatrix} \mathbf{I} & 0 & 0 \\ -\mathbf{B}_{22}^{-1}\mathbf{B}_{21} & \mathbf{B}_{22}^{-1} & 0 \\ * & * & \mathbf{I} \end{bmatrix} \begin{matrix} \leftarrow & D \text{ row} \\ \leftarrow & D(D-1)/2 \text{ row} \\ \leftarrow & N_v - D(D+1)/2 \text{ row} \end{matrix} \tag{3.56}
$$

$$
\begin{matrix} \uparrow & \uparrow & \uparrow \\ D \text{ col} & \frac{D(D-1)}{2} \text{ col} & N_v - \frac{D(D+1)}{2} \text{ col} \end{matrix}
$$

Let

$$
\hat{\mathbf{B}} = \text{diag}\{\mathbf{I}, \ \mathbf{B}_{22}, \ \mathbf{I}\}, \tag{3.57}
$$

be the diagonal blocks of $\mathbf{B}$. The inverse of $\hat{\mathbf{B}}$ is $\hat{\mathbf{B}}^{-1} = \text{diag}\{\mathbf{I}, \ \mathbf{B}_{22}^{-1}, \ \mathbf{I}\}$, and we have

$$
\hat{\mathbf{B}}\mathbf{B}^{-1} = \begin{bmatrix} \mathbf{I} & 0 & 0 \\ -\mathbf{B}_{21} & \mathbf{I} & 0 \\ * & * & \mathbf{I} \end{bmatrix}. \tag{3.58}
$$

Since $\mathbf{B}$ is nonsingular, for an arbitrary $j \in \{1, \cdots, N_v\}$, we let

$$
\boldsymbol{v}^{(j)} = \hat{\mathbf{B}}\mathbf{B}^{-1}I_j, \tag{3.59}
$$

where I_j is the j-th column of the $N_v \times N_v$ identity matrix. Actually, $\boldsymbol{v}^{(j)}$ is the j-th column of $\hat{\mathbf{B}}\mathbf{B}^{-1}$. Notice in (3.50) for any α that $|\alpha| = M$, $\mathcal{N}_{D-1}(\hat{\alpha})$-th component of $\boldsymbol{z}_\lambda$ is $He_{\alpha_1+1}(\lambda)$. For the α satisfying

$$
|\alpha| = M \text{ and } j = \mathcal{N}_{D-1}(\hat{\alpha}), \ k = \alpha_1 + 1,
$$

we choose λ such that
$$He_k(\lambda) = 0.$$

Then we have that the j-th component of $\boldsymbol{z}_\lambda$ vanishes
$$z_{\lambda,j} = He_k(\lambda) = 0.$$

This makes (3.51) valid that
$$\mathbf{B}\hat{\mathbf{B}}^{-1}\boldsymbol{v}^{(j)} = \rho I_j, \quad z_{\lambda,j} = 0, \quad j = 1, \cdots, N_v. \tag{3.60}$$

Since $\boldsymbol{r}$ is depended only on $\boldsymbol{v}$ and λ, we denote $\boldsymbol{r}_{\hat{\alpha},i}$ to be the vector prescribed by the given $\boldsymbol{v}^{(j)}$, $j = \mathcal{N}_{D-1}(\hat{\alpha})$, and $\lambda = C_{i,k}$, when $k = \alpha_1 + 1$, for arbitrary $|\alpha| = M$, $i = 1$, $\cdots$, k. It is clear that $C_{i,k}$ and $\boldsymbol{r}_{\hat{\alpha},i}$ are a pair of eigenvalue and eigenvector of $\tilde{\mathbf{A}}_M$ that

$$\tilde{\mathbf{A}}_M \boldsymbol{r}_{\hat{\alpha},i} = C_{i,k} \boldsymbol{r}_{\hat{\alpha},i}.$$

The eigenvectors of $\tilde{\mathbf{A}}_M$ can be divided into a cluster of classes, each of which is: for arbitrary $|\alpha| = M$,
$$\{\boldsymbol{r}_{\hat{\alpha},i} \mid i = 1, \cdots, k, \ k = \alpha_1 + 1\}.$$

This fact essentially stems from the reducibility of the matrix $\tilde{\mathbf{A}}_M$.

Notice that

1. The components of $\boldsymbol{v}$ are a subset of $\boldsymbol{r}$'s components, linearly independent $\boldsymbol{v}^{(j)}$'s determine linearly independent $\boldsymbol{r}$'s;

2. Eigenvectors belongs to different eigenvalues are orthogonal and the k zeros of Hermite polynomial $He_k(\lambda)$ are different.

We have that $\boldsymbol{r}_{\hat{\alpha},i}$, $i = 1, \cdots, k$, when $|\alpha| = M$ and $k = \alpha_1 + 1$ are linearly independent and the matrix $\tilde{\mathbf{A}}_M$ has

$$\sum_{k=1}^{M+1} k \binom{D-1+M-k}{M+1-k} = \binom{M+D}{D} = N \tag{3.61}$$

linearly independent eigenvectors.

On the other hand, for arbitrary $\boldsymbol{r}_{\hat{\alpha},i}$, there exists a unique β satisfying $\beta = (i-1)e_1 + \tilde{\alpha}$, hence, there is a one-one mapping between $\boldsymbol{r}_{\hat{\alpha},i}$ and α with $|\alpha| \leq M$. So we can also get $\tilde{\mathbf{A}}_M$ has N linearly independent eigenvectors. This completes the proof. $\square$

With the help of Lemma 2, it is not difficult to get the following result:

Lemma 3. *Let*

$$\tilde{\mathcal{P}}_{1,m} = He_{m+1}(\lambda), \quad m \in \mathbb{N}, \tag{3.62}$$

$$\tilde{\mathcal{P}}_{D,m} = \prod_{k=0}^{m} \tilde{\mathcal{P}}_{D-1,k}, \quad 1 < D \in \mathbb{N}^+. \tag{3.63}$$

$\tilde{\mathcal{P}}_{D,M}$ *is the characteristic polynomial of* $\tilde{\mathbf{A}}_M$.

Proof. In case of $D = 1$, the result has been proved in [1]. Here we give the proof for $D \geq 2$. In the proof of lemma 2, (3.60) shows that the characteristic polynomial of $\tilde{\mathbf{A}}_M$ is

$$\tilde{\mathcal{P}}_{D,M} = \prod_{k=1}^{M+1} He_k(\lambda)^{\binom{D-1+M-k}{M+1-k}} = \prod_{k=1}^{M+1} He_k(\lambda)^{\binom{D-1+M-k}{D-2}}. \tag{3.64}$$

Now we need only to prove that $\tilde{\mathcal{P}}_{D,M}$ satisfies (3.63). Here we use induction argument on D. As $D = 2$, (3.64) can be written as

$$\tilde{\mathcal{P}}_{2,M} = \prod_{k=1}^{M+1} He_k(\lambda) = \prod_{m=0}^{M} \tilde{\mathcal{P}}_{1,m}. \tag{3.65}$$

We assume that (3.63) holds for $D - 1, D \geq 3$. With the induction hypothesis, we have

$$\begin{aligned}
\prod_{m=0}^{M} \tilde{\mathcal{P}}_{D-1,m} &= \prod_{m=0}^{M} \left(\prod_{k=1}^{m+1} He_k(\lambda)^{\binom{D-1-1+m-k}{D-1-2}} \right) \\
&= \prod_{k=1}^{M+1} He_k(\lambda)^{\sum_{m=k-1}^{M} \binom{D-2+m-k}{D-3}} \\
&= \prod_{k=1}^{M+1} He_k(\lambda)^{\binom{D-1+M-k}{D-2}} \\
&= \tilde{\mathcal{P}}_{D,M}.
\end{aligned} \tag{3.66}$$

This completes the proof. □

With the relation of $\tilde{\mathbf{A}}_M$ and $\hat{\mathbf{A}}_M$ (3.22), we have the following theorem.

Theorem 2. *Let*

$$\mathcal{P}_{1,m} = He_{m+1}\left(\frac{\lambda - u_1}{\sqrt{\theta}}\right) \theta^{(m+1)/2}, \quad m \in \mathbb{N}, \tag{3.67}$$

$$\mathcal{P}_{D,m} = \prod_{k=0}^{m} \mathcal{P}_{D-1,k}, \quad 1 < D \in \mathbb{N}^+. \tag{3.68}$$

$\mathcal{P}_{D,M}$ *is the characteristic polynomial of* $\hat{\mathbf{A}}_M$. *And* $\hat{\mathbf{A}}_M$ *has* N *linearly independent eigenvectors, which read*

$$\hat{\boldsymbol{r}}_{\hat{\alpha},i} = \boldsymbol{\Lambda}^{-1}\boldsymbol{r}_{\hat{\alpha},i}, \text{ for eigenvalue } \lambda_{i,k} = u_1 + \mathrm{C}_{i,k}\sqrt{\theta} \tag{3.69}$$

for all $|\alpha| = M$, $i = 1, \cdots, k$, *whereas* $k = \alpha_1 + 1$.

Proof. Since

$$\hat{\mathbf{A}}_M = u_1\mathbf{I} + \sqrt{\theta}\boldsymbol{\Lambda}^{-1}\tilde{\mathbf{A}}_M\boldsymbol{\Lambda}, \tag{3.70}$$

and $\boldsymbol{\Lambda}$ is nonsingular, so any $\boldsymbol{r}_{\hat{\alpha},i} \in \mathbb{R}^N$ is the eigenvector of $\tilde{\mathbf{A}}_M$ for the eigenvalue $\mathrm{C}_{i,k}$, then $\hat{\boldsymbol{r}}_{\hat{\alpha},i} = \boldsymbol{\Lambda}^{-1}\boldsymbol{r}_{\hat{\alpha},i}$ is the eigenvector of $\hat{\mathbf{A}}_M$ for the eigenvalue $u_1 + \mathrm{C}_{i,k}\sqrt{\theta}$. Using Lemma 1 and discussion in Lemma 2, we obtain (3.69). Lemma 2 shows $\tilde{\mathbf{A}}_M$ has N linearly independent eigenvectors, so $\hat{\mathbf{A}}_M$ also has N linearly independent eigenvectors and (3.69) gives a set of basis.

Lemma 2 and (3.69) show that the characteristic polynomial of $\hat{\mathbf{A}}_M$ is

$$\mathcal{P}_{D,M} = \prod_{k=1}^{M+1} \left(He_k \left(\frac{\lambda - u_1}{\sqrt{\theta}} \right) \theta^{k/2} \right)^{\binom{D-1+M-k}{D-2}}. \tag{3.71}$$

Similar as that in the proof of Lemma 3, $\mathcal{P}_{D,M}$ is thus the characteristic polynomial of $\hat{\mathbf{A}}_M$. $\qquad\square$

Theorem 1 is now straightforward:

Proof of Theorem 1. With Theorem 2, we declare that $\hat{\mathbf{A}}_M$ is diagonalisable with real eigenvalues directly, that is, the moment system (3.20) is hyperbolic. $\qquad\square$

4   System in Multi-dimensional Spatial Space

As the main result of this paper, here we give the general hyperbolic moment system containing all moments with orders lower than M. Without the assumption that the dependence of f on $x_2, \cdots, x_D$ is homogeneous, according to the discussions in Section 2.3, Grad's moment system can be written in the following form:

$$\frac{\partial \boldsymbol{w}}{\partial t} + \sum_{j=1}^{D} \mathbf{M}_j(\boldsymbol{w}) \frac{\partial \boldsymbol{w}}{\partial x_j} = 0, \tag{4.1}$$

where $\boldsymbol{w}$ remains the same definition as the one-dimensional case (3.9), and $\mathbf{M}_j$, $j = 1, \cdots, D$ are square matrices depending on $\boldsymbol{w}$. Comparing with (3.10), one immediately has $\mathbf{M}_1 = \mathbf{A}_M$. Similar as Definition 2, we give the following definition:

Definition 3. *For $j = 1, \cdots, D$, $\hat{\mathbf{M}}_j$ is called the regularized matrix of the matrix $\mathbf{M}_j$, if it satisfies that for any admissible $\boldsymbol{w}$,*

$$\hat{\mathbf{M}}_j \frac{\partial \boldsymbol{w}}{\partial x_j} = \mathbf{M}_j \frac{\partial \boldsymbol{w}}{\partial x_j} - \sum_{|\alpha|=M} \mathcal{R}_{M,D}^j(\alpha) I_{\mathcal{N}_D(\alpha)}, \tag{4.2}$$

where I_k is the k-th column of the $N \times N$ identity matrix.

Now the multi-dimensional regularized moment equations can be written as

$$\frac{\partial \boldsymbol{w}}{\partial t} + \sum_{j=1}^{D} \hat{\mathbf{M}}_j(\boldsymbol{w}) \frac{\partial \boldsymbol{w}}{\partial x_j} = 0. \tag{4.3}$$

Recalling that

$$\mathcal{R}_{M,D}(\alpha) = \sum_{j=1}^{D} \mathcal{R}_{M,D}^j(\alpha), \tag{4.4}$$

one finds that the multi-dimensional regularized moment system is obtained by subtracting $\mathcal{R}_{M,D}(\alpha)$ from (4.2) for all $|\alpha| = M$. Applying such an operation on (2.7), we can

reformulate the regularized moment system as

$$
\begin{aligned}
&\left(\frac{\partial f_\alpha}{\partial t} + \sum_{d=1}^{D} \frac{\partial u_d}{\partial t} f_{\alpha - e_d} + \frac{1}{2} \frac{\partial \theta}{\partial t} \sum_{d=1}^{D} f_{\alpha - 2e_d} \right) \\
&\quad + \sum_{j=1}^{D} \left(\theta \frac{\partial f_{\alpha - e_j}}{\partial x_j} + u_j \frac{\partial f_\alpha}{\partial x_j} + (1 - \delta_{|\alpha|,M})(\alpha_j + 1) \frac{\partial f_{\alpha + e_j}}{\partial x_j} \right) \\
&\quad + \sum_{j=1}^{D} \sum_{d=1}^{D} \frac{\partial u_d}{\partial x_j} \left(\theta f_{\alpha - e_d - e_j} + u_j f_{\alpha - e_d} + (1 - \delta_{|\alpha|,M})(\alpha_j + 1) f_{\alpha - e_d + e_j} \right) \\
&\quad + \frac{1}{2} \sum_{j=1}^{D} \sum_{d=1}^{D} \frac{\partial \theta}{\partial x_j} \left(\theta f_{\alpha - 2e_d - e_j} + u_j f_{\alpha - 2e_d} + (1 - \delta_{|\alpha|,M})(\alpha_j + 1) f_{\alpha - 2e_d + e_j} \right) = 0,
\end{aligned}
$$

$$
|\alpha| \leqslant M. \tag{4.5}
$$

Actually, (4.5) is away from (4.3) only by a linear transformation due to (2.8) to eliminate the time derivatives of u_d and θ. Precisely speaking, there exists an invertible matrix $\mathbf{T}(\boldsymbol{w})$ depending on $\boldsymbol{w}$ such that (4.5) is identical to the following system:

$$
\mathbf{T}(\boldsymbol{w}) \frac{\partial \boldsymbol{w}}{\partial t} + \sum_{j=1}^{D} \mathbf{T}(\boldsymbol{w}) \hat{\mathbf{M}}_j(\boldsymbol{w}) \frac{\partial \boldsymbol{w}}{\partial x_j} = 0, \tag{4.6}
$$

If we let all partial derivatives with respect to x_j with $j > 1$ to be zero, (4.5) reduces to the one-dimensional hyperbolic moment system (3.20) in Section 3. Comparison of (4.3) and (3.20) clearly shows that $\hat{\mathbf{M}}_1 = \hat{\mathbf{A}}_M$.

The following theorem declares the hyperbolicity[2] of the multi-dimensional regularized moment system (4.3):

Theorem 3. *The regularized moment system* (4.3) *is hyperbolic for any admissible $\boldsymbol{w}$. Precisely, for a given unit vector $\boldsymbol{n} = (n_1, \cdots, n_D)$, there exists a constant matrix $\mathbf{R}$ partially depending on $\boldsymbol{n}$ that*

$$
\sum_{j=1}^{D} n_j \hat{\mathbf{M}}_j(\boldsymbol{w}) = \mathbf{R}^{-1} \hat{\mathbf{A}}_M(\mathbf{R}\boldsymbol{w})\mathbf{R}, \tag{4.7}
$$

and this matrix is diagonalizable with eigenvalues as

$$
\boldsymbol{u} \cdot \boldsymbol{n} + \mathrm{C}_{n,m}\sqrt{\theta}, \qquad 1 \leqslant n \leqslant m \leqslant M + 1. \tag{4.8}
$$

Actually, this theorem gives the rotation invariance of the regularized moment system and its globally hyperbolicity. Since the translation invariance of the system is apparent, it is concluded that the regularized system is Galilean invariant. Precisely, if another coordinates $(\tilde{x}_1, \ldots, \tilde{x}_D)$ are chosen and the vector $\boldsymbol{n}$ is along the $\tilde{x}_1$-axis, then the rotated moment system is equivalent to the original one. This result is easy to understand: on one hand, Grad's moment system is rotationally invariant, since the full M-degree polynomials are used in the truncated Hermite expansion; on the other hand, our regularization

---

[2]For multi-dimensional quasi-linear systems, we refer the readers to [8] for the definition of hyperbolicity.

is symmetric in every direction, which can be considered as "isotropic" in some sense. However, a rigorous proof of this theorem is rather tedious.

In the literature, two types of indices have been used in the moment methods. In Grad's paper [7], indices such as

$$\vartheta = (\vartheta_1, \cdots, \vartheta_m) \in \mathcal{D}^m \tag{4.9}$$

is used to denote the m-th order moments, while in [2], the symbols

$$\alpha = (\alpha_1, \cdots, \alpha_D) \in \mathbb{N}^D \tag{4.10}$$

is used as the subscripts of $|\alpha|$-th order moments. The former is convenient for mathematical proofs, while the latter is easier to use in the numerical implementation, since for (4.10), the map from the index set to the moment set is a bijection, while this is not true for (4.9). If (4.10) and (4.9) represent the same moment, then one has

$$\alpha = e_{\vartheta_1} + \cdots + e_{\vartheta_m}, \quad m = |\alpha|. \tag{4.11}$$

Below, both types of indices are needed in the proof of rotation invariance, and we will always use the variant forms of Greek letters such as ϑ and φ to denote the Grad-type indices, and normal Greek letters such as α and β will be used to denote indices like (4.10). The Greek letter "sigma" denotes the conversion between them. Supposing (4.11) holds, we write

$$\alpha = \sigma(\vartheta), \quad \vartheta = \varsigma(\alpha). \tag{4.12}$$

That is, the normal form of sigma $\sigma(\cdot)$ converts indices like (4.9) to indices like (4.10), and the variant form of sigma $\varsigma(\cdot)$ does the inverse conversion. Note that for a given α, the Grad-type index ϑ satisfying (4.11) is not uniquely determined. Define

$$\Sigma(\alpha) = \{\vartheta \in \mathcal{D}^{|\alpha|} \mid \sigma(\vartheta) = \alpha\}, \tag{4.13}$$

and then in most cases, $\Sigma(\alpha)$ has more than one element. For example, if $D = 2$ and $\alpha = (2, 2)$, then

$$\Sigma(\alpha) = \{(1, 1, 2, 2), (1, 2, 1, 2), (1, 2, 2, 1), (2, 2, 1, 1), (2, 1, 2, 1), (2, 1, 1, 2)\}. \tag{4.14}$$

Thus $\varsigma(\alpha)$ has multiple values. However, there is always one special element $\vartheta \in \Sigma(\alpha)$ satisfying

$$\vartheta_1 \leqslant \cdots \leqslant \vartheta_{|\alpha|}, \tag{4.15}$$

and we use this element as the value of $\varsigma(\alpha)$. It is easy to find

$$\sigma(\varsigma(\alpha)) = \alpha. \tag{4.16}$$

Additionally, we use $\sigma_i(\vartheta)$ to denote the i-th component of $\sigma(\vartheta)$.

Based on these symbols, we have the following lemma:

Lemma 4. *Suppose $\alpha \in \mathbb{N}^D$ and $F(\cdot)$ is a function on $\mathcal{D}^m$. If F satisfies that $F(\varphi)$ is zero when $\sigma_i(\varphi) < \alpha_i$ for some $i \in \mathcal{D}$, then the following equality holds:*

$$\sum_{\varphi \in \mathcal{D}^m} F(\varphi) = \sum_{\substack{\beta \in \mathbb{N}^D \\ |\beta| = m - |\alpha|}} \sum_{\varphi \in \sum(\beta + \alpha)} F(\varphi). \tag{4.17}$$

23

Proof. It is obvious that

$$\mathcal{I} \triangleq \bigcup_{\substack{\beta \in \mathbb{N}^D \\ |\beta| = m - |\alpha|}} \Sigma(\beta + \alpha) \subset \mathcal{D}^m, \tag{4.18}$$

and there are no duplicate elements in the union since $\Sigma(\beta + \alpha) \cap \Sigma(\tilde{\beta} + \alpha) = \emptyset$ if $\beta \neq \tilde{\beta}$. Thus it only remains to prove that $\varphi \in \mathcal{I}$ if

$$\varphi \in \mathcal{D}^m, \quad \text{and} \quad \sigma_i(\varphi) \geqslant \alpha_i, \quad \forall i \in \mathcal{D}. \tag{4.19}$$

This is true since $\varphi \in \Sigma(\beta + \alpha)$ for $\beta = \sigma(\varphi) - \alpha$. $\qquad\square$

As a special case of Lemma 4, we set $\alpha = 0$ and have

$$\sum_{\varphi \in \mathcal{D}^m} F(\varphi) = \sum_{\substack{\beta \in \mathbb{N}^D \\ |\beta| = m}} \sum_{\varphi \in \ (\beta)} F(\varphi). \tag{4.20}$$

Here $F(\cdot)$ is an arbitrary function on $\mathcal{D}^m$.

Some more symbols are introduced as follows. All m-permutations of the set $\{1, \cdots, n\}$ form the following set:

$$\mathcal{A}_n^m = \{\varpi = (\varpi_1, \cdots, \varpi_m) \in \{1, \cdots, n\}^m \mid \varpi_i \neq \varpi_j \text{ if } i \neq j\}, \quad \forall m, n \in \mathbb{N}, \quad n \geqslant m, \tag{4.21}$$

which contains $n!/m!$ elements. Thus when we want to construct a short vector using the components of a long vector, we will use the following notation:

$$\vartheta_\varpi = (\vartheta_{\varpi_1}, \cdots, \vartheta_{\varpi_m}) \in \mathcal{D}^m, \quad \forall \vartheta \in \mathcal{D}^n, \quad \varpi \in \mathcal{A}_n^m. \tag{4.22}$$

The remaining part is denoted as $\vartheta \backslash \vartheta_\varpi$. For example, if $\vartheta = (1, 3, 2, 3, 1, 2, 1)$ and $\varpi = (5, 2, 4)$, then

$$\vartheta_\varpi = (1, 3, 3), \quad \vartheta \backslash \vartheta_\varpi = (1, 2, 2, 1). \tag{4.23}$$

Below, $\mathbf{G} = (g_{ij})_{D \times D}$ stands for the rotation matrix, and we suppose $\mathbf{G}$ is orthogonal and its the determinant is 1. Define

$$\Pi_g(\vartheta, \varphi) = \prod_{i=1}^n g_{\vartheta_i \varphi_i}, \quad \forall \vartheta, \varphi \in \mathcal{D}^n, \tag{4.24}$$

and then we have the following lemma:

Lemma 5. *For a given matrix* $\mathbf{G}$ *and multi-indices* $\alpha, \beta \in \mathbb{N}^D$, *the following equality holds for arbitrary* $\vartheta \in \mathcal{D}^{|\alpha|+|\beta|}$:

$$\sum_{\varphi \in \ (\alpha+\beta)} \frac{\sigma(\varphi)!}{\sigma(\vartheta)!} \Pi_g(\vartheta, \varphi) = \frac{\alpha!}{\sigma(\vartheta)!} \sum_{\varpi \in \mathcal{A}_{|\alpha|+|\beta|}^{|\beta|}} \Pi_g(\vartheta_\varpi, \varsigma(\beta)) \sum_{\varphi \in \ (\alpha)} \Pi_g(\vartheta \backslash \vartheta_\varpi, \varphi). \tag{4.25}$$

Proof. We first consider the case $|\beta| = 1$. Suppose $\beta = e_d$, and then (4.25) becomes

$$\sum_{\varphi \in \ (\alpha+e_d)} \frac{\sigma(\varphi)!}{\sigma(\vartheta)!} \Pi_g(\vartheta, \varphi) = \frac{\alpha!}{\sigma(\vartheta)!} \sum_{i=1}^{|\alpha|+1} g_{\vartheta_i d} \sum_{\varphi \in \ (\alpha)} \Pi_g(\vartheta \backslash \vartheta_i, \varphi). \tag{4.26}$$

24

For $\varphi \in \Sigma(\alpha + e_d)$, one has $\sigma(\varphi)! = (\alpha_d + 1)\alpha!$. Thus (4.26) is equivalent to

$$(\alpha_d + 1) \sum_{\varphi \in \Sigma(\alpha + e_d)} \Pi_g(\vartheta, \varphi) = \sum_{i=1}^{|\alpha|+1} g_{\vartheta_i d} \sum_{\varphi \in \Sigma(\alpha)} \Pi_g(\vartheta \backslash \vartheta_i, \varphi). \tag{4.27}$$

For an arbitrary $\varphi \in \Sigma(\alpha + e_d)$, if $\varphi_i = d$, then $\Pi_g(\vartheta, \varphi) = g_{\vartheta_i d}\Pi_g(\vartheta \backslash \vartheta_i, \varphi \backslash \varphi_i)$, and $\varphi \backslash \varphi_i \in \Sigma(\alpha)$. Since there are $(\alpha_d + 1)$ choices of i such that $\varphi_i = d$, the product $\Pi_g(\vartheta, \varphi)$ appears $(\alpha_d + 1)$ times in the right hand side of (4.27). This proves (4.26).

Suppose the lemma holds for $|\beta| = m - 1$, and we are going to prove the case $|\beta| = m$. In order to use the technique of induction, we choose $d \in \{1, \cdots, D\}$ such that $\beta_d > 0$, and let $\beta' = \beta - e_d$. Thus $|\beta'| = m - 1$. Applying (4.26), one has

$$\sum_{\varphi \in \Sigma(\alpha+\beta)} \frac{\sigma(\varphi)!}{\sigma(\vartheta)!} \Pi_g(\vartheta, \varphi) = \sum_{\varphi \in \Sigma(\alpha+\beta'+e_d)} \frac{\sigma(\varphi)!}{\sigma(\vartheta)!} \Pi_g(\vartheta, \varphi)$$

$$= \frac{(\alpha + \beta')!}{\sigma(\vartheta)!} \sum_{i=1}^{|\alpha|+|\beta|} g_{\vartheta_i d} \sum_{\varphi \in \Sigma(\alpha+\beta')} \Pi_g(\vartheta \backslash \vartheta_i, \varphi) \tag{4.28}$$

$$= \sum_{i=1}^{|\alpha|+|\beta|} \frac{\sigma(\vartheta \backslash \vartheta_i)!}{\sigma(\vartheta)!} g_{\vartheta_i d} \sum_{\varphi \in \Sigma(\alpha+\beta')} \frac{\sigma(\varphi)!}{\sigma(\vartheta \backslash \vartheta_i)!} \Pi_g(\vartheta \backslash \vartheta_i, \varphi)$$

Defining $\hat{\vartheta}^i = \vartheta \backslash \vartheta_i$, and using the inductive assumption, one obtains

$$\sum_{\varphi \in \Sigma(\alpha+\beta)} \frac{\sigma(\varphi)!}{\sigma(\vartheta)!} \Pi_g(\vartheta, \varphi) = \frac{\alpha!}{\sigma(\vartheta)!} \sum_{i=1}^{|\alpha|+|\beta|} g_{\vartheta_i d} \sum_{\varpi \in \mathcal{A}_{|\alpha|+|\beta|-1}^{|\beta|-1}} \Pi_g(\hat{\vartheta}_\varpi^i, \sigma(\beta')) \sum_{\varphi \in \Sigma(\alpha)} \Pi_g(\hat{\vartheta}^i \backslash \hat{\vartheta}_\varpi^i, \varphi) \tag{4.29}$$

It is evident that the right hand sides of (4.25) and (4.29) are the same. Thus the lemma is proved. $\qquad\square$

Now let us start the rotation. We first define

$$\tilde{x}_i = \sum_{j=1}^{D} g_{ij} x_j, \quad i = 1, \cdots, D, \tag{4.30}$$

and denote by $\tilde{\rho}$, $\tilde{\boldsymbol{u}}$, $\tilde{\theta}$ the density, macroscopic velocity and temperature in the new coordinates $\tilde{\boldsymbol{x}} = (\tilde{x}_1, \cdots, \tilde{x}_D)$. If we define $\tilde{\boldsymbol{\xi}} = \mathbf{G}\boldsymbol{\xi}$, then the orthogonality of $\mathbf{G}$ shows

$$\tilde{\rho} = \int_{\mathbb{R}^D} f(\boldsymbol{\xi}) \, d\tilde{\boldsymbol{\xi}} = \int_{\mathbb{R}^D} f(\boldsymbol{\xi}) \, d\boldsymbol{\xi} = \rho,$$

$$\tilde{\rho}\tilde{\boldsymbol{u}} = \int_{\mathbb{R}^D} \tilde{\boldsymbol{\xi}} f(\boldsymbol{\xi}) \, d\tilde{\boldsymbol{\xi}} = \int_{\mathbb{R}^D} \mathbf{G}\boldsymbol{\xi} f(\boldsymbol{\xi}) \, d\boldsymbol{\xi} = \rho\mathbf{G}\boldsymbol{u}, \tag{4.31}$$

$$\tilde{\rho}\tilde{\theta} = \frac{1}{D} \int_{\mathbb{R}^D} |\tilde{\boldsymbol{\xi}} - \tilde{\boldsymbol{u}}|^2 f(\boldsymbol{\xi}) \, d\tilde{\boldsymbol{\xi}} = \frac{1}{D} \int_{\mathbb{R}^D} |\boldsymbol{\xi} - \boldsymbol{u}|^2 f(\boldsymbol{\xi}) \, d\boldsymbol{\xi} = \rho\theta,$$

and it follows immediately that

$$\tilde{\theta} = \theta, \qquad \tilde{u}_i = \sum_{j=1}^{D} g_{ij} u_j, \quad i = 1, \cdots, D. \tag{4.32}$$

25

Now we consider the general moments $\tilde{f}_\alpha$ in the coordinates $\tilde{\boldsymbol{x}}$. Define $\boldsymbol{z} = (\boldsymbol{\xi} - \boldsymbol{u})/\sqrt{\theta}$ and $\tilde{\boldsymbol{z}} = (\tilde{\boldsymbol{\xi}} - \tilde{\boldsymbol{u}})/\sqrt{\tilde{\theta}}$. Then $\tilde{\boldsymbol{z}} = \mathbf{G}\boldsymbol{z}$. The orthogonality of Hermite polynomials gives

$$
\begin{aligned}
f_\alpha &= \frac{(2\pi)^D \theta^{|\alpha|+D}}{\alpha!} \int_{\mathbb{R}^D} f(\boldsymbol{u} + \sqrt{\theta}\boldsymbol{z}) \mathcal{H}_{\theta,\alpha}(\boldsymbol{z}) \exp\left(-\frac{|\boldsymbol{z}|^2}{2}\right) \mathrm{d}\boldsymbol{z}, \\
\tilde{f}_\alpha &= \frac{(2\pi)^D \tilde{\theta}^{|\alpha|+D}}{\alpha!} \int_{\mathbb{R}^D} f(\boldsymbol{u} + \sqrt{\theta}\boldsymbol{z}) \mathcal{H}_{\theta,\alpha}(\tilde{\boldsymbol{z}}) \exp\left(-\frac{|\tilde{\boldsymbol{z}}|^2}{2}\right) \mathrm{d}\tilde{\boldsymbol{z}},
\end{aligned}
\tag{4.33}
$$

From the definition of Hermite polynomials (2.4), it is easy to find that (2.3) can be rewritten as

$$
\mathcal{H}_{\theta,\alpha}(\boldsymbol{z}) = (-1)^{|\alpha|} (2\pi)^{-\frac{D}{2}} \theta^{-\frac{D+|\alpha|}{2}} \frac{\partial^{|\alpha|}}{\partial^\alpha \boldsymbol{z}} \exp\left(-\frac{|\boldsymbol{z}|^2}{2}\right).
\tag{4.34}
$$

Applying the chain rule of differentiation, we obtain

$$
\begin{aligned}
\mathcal{H}_{\theta,\alpha}(\tilde{\boldsymbol{z}}) &= (-1)^{|\alpha|} (2\pi)^{-\frac{D}{2}} \theta^{-\frac{D+|\alpha|}{2}} \sum_{\varphi \in \mathcal{D}^{|\alpha|}} \Pi_g(\varsigma(\alpha), \varphi) \frac{\partial^{|\alpha|}}{\partial^{\sigma(\varphi)} \boldsymbol{z}} \exp\left(-\frac{|\boldsymbol{z}|^2}{2}\right) \\
&= \sum_{\varphi \in \mathcal{D}^{|\alpha|}} \Pi_g(\varsigma(\alpha), \varphi) \mathcal{H}_{\theta,\sigma(\varphi)}(\boldsymbol{z}).
\end{aligned}
\tag{4.35}
$$

Collecting (4.33) and (4.35), one has

$$
\begin{aligned}
\tilde{f}_\alpha &= \frac{(2\pi)^D \theta^{|\alpha|+D}}{\alpha!} \sum_{\varphi \in \mathcal{D}^{|\alpha|}} \Pi_g(\varsigma(\alpha), \varphi) \int_{\mathbb{R}^D} f(\boldsymbol{u} + \sqrt{\theta}\boldsymbol{z}) \mathcal{H}_{\theta,\sigma(\varphi)}(\boldsymbol{z}) \exp\left(-\frac{|\boldsymbol{z}|^2}{2}\right) \mathrm{d}\boldsymbol{z} \\
&= \sum_{\varphi \in \mathcal{D}^{|\alpha|}} \frac{\sigma(\varphi)!}{\alpha!} \Pi_g(\varsigma(\alpha), \varphi) f_{\sigma(\varphi)}.
\end{aligned}
\tag{4.36}
$$

As in (3.9), all the rotated moments can also be collected into a vector denoted as $\tilde{\boldsymbol{w}}$. The equations (4.36) and (4.32) directly give the following result:

Lemma 6. *Based on the expressions of the rotated moments* (4.36) *and* (4.32), *the following equalities hold for arbitrary* $\alpha \in \mathbb{N}^D$:

$$
\sum_{d=1}^D \tilde{u}_d \frac{\partial \tilde{f}_\alpha}{\partial \tilde{x}_d} = \sum_{d=1}^D \sum_{\varphi \in \mathcal{D}^{|\alpha|}} \frac{\sigma(\varphi)!}{\alpha!} \Pi_g(\varsigma(\alpha), \varphi) \cdot u_d \frac{\partial f_{\sigma(\varphi)}}{\partial x_d},
\tag{4.37}
$$

$$
\sum_{d=1}^D (\alpha_d + 1) \frac{\partial \tilde{f}_{\alpha+e_d}}{\partial \tilde{x}_d} = \sum_{d=1}^D \sum_{\varphi \in \mathcal{D}^{|\alpha|}} \frac{\sigma(\varphi)!}{\alpha!} \Pi_g(\varsigma(\alpha), \varphi) \cdot (\sigma_d(\varphi) + 1) \frac{\partial f_{\sigma(\varphi)+e_d}}{\partial x_d},
\tag{4.38}
$$

where $\sigma_d(\varphi)$ *is the d-th component of* $\sigma(\varphi)$.

Proof. Using (4.36) and (4.32) directly, we get

$$
\sum_{d=1}^D \tilde{u}_d \frac{\partial \tilde{f}_\alpha}{\partial \tilde{x}_d} = \sum_{d=1}^D \sum_{\varphi \in \mathcal{D}^{|\alpha|}} \frac{\sigma(\varphi)!}{\alpha!} \Pi_g(\varsigma(\alpha), \varphi) \cdot \sum_{j=1}^D g_{dj} u_j \frac{\partial f_{\sigma(\varphi)}}{\partial \tilde{x}_d}.
\tag{4.39}
$$

Equation (4.30) shows that

$$
\frac{\partial}{\partial x_j} = \sum_{d=1}^D g_{dj} \frac{\partial}{\partial \tilde{x}_d}.
\tag{4.40}
$$

Thus (4.37) is the direct result of (4.39) and (4.40).

The proof of (4.38) is also straightforward:

$$\sum_{d=1}^{D}(\alpha_d+1)\frac{\partial \tilde{f}_{\alpha+e_d}}{\partial \tilde{x}_d} = \sum_{d=1}^{D}(\alpha_d+1)\sum_{\varphi\in\mathcal{D}^{|\alpha|+1}}\frac{\sigma(\varphi)!}{(\alpha+e_d)!}\Pi_g(\varsigma(\alpha+e_d),\varphi)\frac{\partial f_{\sigma(\varphi)}}{\partial \tilde{x}_d}.$$

$$= \sum_{d=1}^{D}\sum_{j=1}^{D}g_{dj}\sum_{\varphi\in\mathcal{D}^{|\alpha|}}\frac{(\sigma(\varphi)+e_j)!}{\alpha!}\Pi_g(\varsigma(\alpha),\varphi)\frac{\partial f_{\sigma(\varphi)+e_j}}{\partial \tilde{x}_d} \qquad (4.41)$$

$$= \sum_{j=1}^{D}\sum_{\varphi\in\mathcal{D}^{|\alpha|}}\frac{\sigma(\varphi)!}{\alpha!}\Pi_g(\varsigma(\alpha),\varphi)\cdot(\sigma_j(\varphi)+1)\frac{\partial f_{\sigma(\varphi)+e_j}}{\partial x_j}.$$

This equality is identical to (4.38). $\qquad\qquad\Box$

Using Lemma 5, it is not difficult to prove the following lemma:

Lemma 7. *The following equalities hold for arbitrary $\alpha \in \mathbb{N}^D$:*

$$\sum_{d=1}^{D}\frac{\partial \tilde{u}_d}{\partial t}\tilde{f}_{\alpha-e_d} = \sum_{d=1}^{D}\sum_{\varphi\in\mathcal{D}^{|\alpha|}}\frac{\sigma(\varphi)!}{\alpha!}\Pi_g(\varsigma(\alpha),\varphi)\frac{\partial u_d}{\partial t}f_{\sigma(\varphi)-e_d}, \qquad (4.42)$$

$$\sum_{d=1}^{D}\frac{\partial \tilde{\theta}}{\partial t}\tilde{f}_{\alpha-2e_d} = \sum_{d=1}^{D}\sum_{\varphi\in\mathcal{D}^{|\alpha|}}\frac{\sigma(\varphi)!}{\alpha!}\Pi_g(\varsigma(\alpha),\varphi)\frac{\partial \theta}{\partial t}f_{\sigma(\varphi)-2e_d}, \qquad (4.43)$$

$$\sum_{d=1}^{D}\frac{\partial \tilde{f}_{\alpha-e_d}}{\partial \tilde{x}_d} = \sum_{d=1}^{D}\sum_{\varphi\in\mathcal{D}^{|\alpha|}}\frac{\sigma(\varphi)!}{\alpha!}\Pi_g(\varsigma(\alpha),\varphi)\frac{\partial f_{\sigma(\varphi)-e_d}}{\partial x_d}, \qquad (4.44)$$

$$\sum_{j=1}^{D}\sum_{d=1}^{D}\frac{\partial \tilde{u}_d}{\partial \tilde{x}_j}\tilde{f}_{\alpha-e_d-e_j} = \sum_{j=1}^{D}\sum_{d=1}^{D}\sum_{\varphi\in\mathcal{D}^{|\alpha|}}\frac{\sigma(\varphi)!}{\alpha!}\Pi_g(\varsigma(\alpha),\varphi)\frac{\partial u_d}{\partial x_j}f_{\sigma(\varphi)-e_d-e_j}, \qquad (4.45)$$

$$\sum_{j=1}^{D}\sum_{d=1}^{D}\frac{\partial \tilde{\theta}}{\partial \tilde{x}_j}\tilde{f}_{\alpha-2e_d-e_j} = \sum_{j=1}^{D}\sum_{d=1}^{D}\sum_{\varphi\in\mathcal{D}^{|\alpha|}}\frac{\sigma(\varphi)!}{\alpha!}\Pi_g(\varsigma(\alpha),\varphi)\frac{\partial \theta}{\partial x_j}f_{\sigma(\varphi)-2e_d-e_j}. \qquad (4.46)$$

Proof. Recalling that $f_\beta = 0$ if β has a negative component and using Lemma 4, we have the following equality:

$$I \triangleq \sum_{d=1}^{D}\sum_{\varphi\in\mathcal{D}^{|\alpha|}}\frac{\sigma(\varphi)!}{\alpha!}\Pi_g(\varsigma(\alpha),\varphi)\frac{\partial u_d}{\partial t}f_{\sigma(\varphi)-e_d}$$

$$= \sum_{d=1}^{D}\sum_{\substack{\beta\in\mathbb{N}^D\\|\beta|=|\alpha|-1}}\sum_{\varphi\in\ (\beta+e_d)}\frac{\sigma(\varphi)!}{\alpha!}\Pi_g(\varsigma(\alpha),\varphi)\frac{\partial u_d}{\partial t}f_\beta. \qquad (4.47)$$

Let $\vartheta = \varsigma(\alpha)$ and use (4.26), and we have

$$I = \sum_{d=1}^{D}\sum_{\substack{\beta\in\mathbb{N}^D\\|\beta|=|\alpha|-1}}\frac{\beta!}{\alpha!}\sum_{i=1}^{|\alpha|}g_{\vartheta_i d}\sum_{\varphi\in\ (\beta)}\Pi_g(\vartheta\backslash\vartheta_i,\varphi)\frac{\partial u_d}{\partial t}f_\beta. \qquad (4.48)$$

27

Now we employ (4.20) to join two of the summation symbols in the above equation:

$$I = \sum_{d=1}^{D} \frac{\partial u_d}{\partial t} \sum_{i=1}^{|\alpha|} g_{\vartheta_i d} \sum_{\varphi \in \mathcal{D}^{|\alpha|-1}} \frac{\sigma(\varphi)!}{\alpha!} \Pi_g(\vartheta \backslash \vartheta_i, \varphi) f_{\sigma(\varphi)}. \tag{4.49}$$

Using (4.36) and (4.32) again, we get

$$I = \sum_{d=1}^{D} \sum_{i=1}^{|\alpha|} \frac{1}{\alpha_{\vartheta_i}} g_{\vartheta_i d} \frac{\partial u_d}{\partial t} \tilde{f}_{\alpha - e_{\vartheta_i}} = \sum_{i=1}^{|\alpha|} \frac{1}{\alpha_{\vartheta_i}} \frac{\partial \tilde{u}_{\vartheta_i}}{\partial t} \tilde{f}_{\alpha - e_{\vartheta_i}} = \sum_{d=1}^{D} \frac{\partial \tilde{u}_d}{\partial t} \tilde{f}_{\alpha - e_d}. \tag{4.50}$$

Thus (4.42) is proved. The equation (4.43) can be proved in a similar way. Setting $\vartheta = \varsigma(\alpha)$ and using Lemma 4, Lemma 5 and (4.36), we obtain

$$
\begin{aligned}
II &\triangleq \sum_{d=1}^{D} \sum_{\varphi \in \mathcal{D}^{|\alpha|}} \frac{\sigma(\varphi)!}{\alpha!} \Pi_g(\varsigma(\alpha), \varphi) \frac{\partial \theta}{\partial t} f_{\sigma(\varphi) - 2e_d} \\
&= \sum_{d=1}^{D} \sum_{\substack{\beta \in \mathbb{N}^D \\ |\beta| = |\alpha| - 2}} \sum_{\varphi \in (\beta + 2e_d)} \frac{\sigma(\varphi)!}{\sigma(\vartheta)!} \Pi_g(\vartheta, \varphi) \frac{\partial \theta}{\partial t} f_\beta \\
&= \sum_{d=1}^{D} \sum_{\substack{\beta \in \mathbb{N}^D \\ |\beta| = |\alpha| - 2}} \frac{\beta!}{\alpha!} \sum_{\substack{i,j=1,\cdots,|\alpha| \\ i \neq j}} g_{\vartheta_i d} g_{\vartheta_j d} \sum_{\varphi \in (\beta)} \Pi_g(\vartheta \backslash \vartheta_{(i,j)}, \varphi) \frac{\partial \theta}{\partial t} f_\beta \\
&= \sum_{d=1}^{D} \sum_{\substack{i,j=1,\cdots,|\alpha| \\ i \neq j}} \frac{(\alpha - e_{\vartheta_i} - e_{\vartheta_j})!}{\alpha!} g_{\vartheta_i d} g_{\vartheta_j d} \frac{\partial \tilde{\theta}}{\partial t} \tilde{f}_{\alpha - e_{\vartheta_i} - e_{\vartheta_j}}.
\end{aligned}
\tag{4.51}
$$

Since $\mathbf{G}$ is an orthogonal matrix, one has

$$\sum_{d=1}^{D} g_{\vartheta_i d} g_{\vartheta_j d} = \delta_{\vartheta_i \vartheta_j}. \tag{4.52}$$

Thus (4.51) can be further simplified as

$$II = \sum_{\substack{i,j=1,\cdots,|\alpha| \\ i \neq j, \vartheta_i = \vartheta_j}} \frac{1}{\alpha_{\vartheta_i}(\alpha_{\vartheta_i} - 1)} \frac{\partial \tilde{\theta}}{\partial t} \tilde{f}_{\alpha - 2e_{\vartheta_i}} = \sum_{d=1}^{D} \frac{\partial \tilde{\theta}}{\partial t} \tilde{f}_{\alpha - 2e_d}, \tag{4.53}$$

which completes the proof of (4.43). The equations (4.44)(4.45)(4.46) can be proved using exactly the same technique. The detailed proofs are omitted here to avoid redundancy. $\square$

It is not difficult to find that (4.42) and (4.43) still hold if we replace t with x_j or $\tilde{x}_j$ for any $j = 1, \cdots, D$. Such observation leads to the following two lemmas:

Lemma 8. *The following equalities hold for arbitrary* $\alpha \in \mathbb{N}^D$:

$$\sum_{j=1}^{D} \sum_{d=1}^{D} \frac{\partial \tilde{u}_d}{\partial \tilde{x}_j} \tilde{u}_j \tilde{f}_{\alpha - e_d} = \sum_{j=1}^{D} \sum_{d=1}^{D} \sum_{\varphi \in \mathcal{D}^{|\alpha|}} \frac{\sigma(\varphi)!}{\alpha!} \Pi_g(\varsigma(\alpha), \varphi) \frac{\partial u_d}{\partial x_j} u_j f_{\sigma(\varphi) - e_d}, \tag{4.54}$$

$$\sum_{j=1}^{D} \sum_{d=1}^{D} \frac{\partial \tilde{\theta}}{\partial \tilde{x}_j} \tilde{u}_j \tilde{f}_{\alpha - 2e_d} = \sum_{j=1}^{D} \sum_{d=1}^{D} \sum_{\varphi \in \mathcal{D}^{|\alpha|}} \frac{\sigma(\varphi)!}{\alpha!} \Pi_g(\varsigma(\alpha), \varphi) \frac{\partial \theta}{\partial x_j} u_j f_{\sigma(\varphi) - 2e_d}. \tag{4.55}$$

Proof. Replacing t with x_j in (4.42), we obtain

$$
\sum_{j=1}^{D}\sum_{d=1}^{D}\sum_{\varphi\in\mathcal{D}^{|\alpha|}}\frac{\sigma(\varphi)!}{\alpha!}\Pi_g(\varsigma(\alpha),\varphi)\frac{\partial u_d}{\partial x_j}u_j f_{\sigma(\varphi)-e_d} = \sum_{j=1}^{D}\sum_{d=1}^{D}\frac{\partial\tilde{u}_d}{\partial x_j}u_j\tilde{f}_{\alpha-e_d}
$$
$$
= \sum_{j=1}^{D}\sum_{d=1}^{D}\sum_{i=1}^{D}g_{ij}\frac{\partial\tilde{u}_d}{\partial\tilde{x}_i}u_j\tilde{f}_{\alpha-e_d} = \sum_{d=1}^{D}\sum_{i=1}^{D}\frac{\partial\tilde{u}_d}{\partial\tilde{x}_i}\tilde{u}_i\tilde{f}_{\alpha-e_d}.
$$

(4.56)

This equation is the same as (4.54). The proof of (4.55) is almost the same. $\square$

Lemma 9. *The following equalities hold for arbitrary $\alpha\in\mathbb{N}^D$:*

$$
\sum_{j=1}^{D}\sum_{d=1}^{D}(\alpha_j+1)\frac{\partial\tilde{u}_d}{\partial\tilde{x}_j}\tilde{f}_{\alpha-e_d+e_j} = \sum_{j=1}^{D}\sum_{d=1}^{D}\sum_{\varphi\in\mathcal{D}^{|\alpha|}}\frac{\sigma(\varphi)!}{\alpha!}\Pi_g(\varsigma(\alpha),\varphi)\cdot(\sigma_j(\varphi)+1)\frac{\partial u_d}{\partial x_j}f_{\sigma(\varphi)-e_d+e_j},
$$
$$
\sum_{j=1}^{D}\sum_{d=1}^{D}(\alpha_j+1)\frac{\partial\tilde{\theta}}{\partial\tilde{x}_j}\tilde{f}_{\alpha-2e_d+e_j} = \sum_{j=1}^{D}\sum_{d=1}^{D}\sum_{\varphi\in\mathcal{D}^{|\alpha|}}\frac{\sigma(\varphi)!}{\alpha!}\Pi_g(\varsigma(\alpha),\varphi)\cdot(\sigma_j(\varphi)+1)\frac{\partial\theta}{\partial x_j}f_{\sigma(\varphi)-2e_d+e_j}.
$$

Proof. Replacing t with $\tilde{x}_j$ and substituting $\alpha+e_j$ for α in (4.42), one has

$$
\sum_{j=1}^{D}\sum_{d=1}^{D}(\alpha_j+1)\frac{\partial\tilde{u}_d}{\partial\tilde{x}_j}\tilde{f}_{\alpha-e_d+e_j}
$$
$$
= \sum_{j=1}^{D}\sum_{d=1}^{D}(\alpha_j+1)\sum_{\varphi\in\mathcal{D}^{|\alpha|+1}}\frac{\sigma(\varphi)!}{(\alpha+e_j)!}\Pi_g(\varsigma(\alpha+e_j),\varphi)\frac{\partial u_d}{\partial\tilde{x}_j}f_{\sigma(\varphi)-e_d}
$$
$$
= \sum_{j=1}^{D}\sum_{d=1}^{D}\sum_{i=1}^{D}g_{ji}\sum_{\varphi\in\mathcal{D}^{|\alpha|}}\frac{(\sigma(\varphi)+e_i)!}{\alpha!}\Pi_g(\varsigma(\alpha),\varphi)\frac{\partial u_d}{\partial\tilde{x}_j}f_{\sigma(\varphi)+e_i-e_d}
$$
$$
= \sum_{d=1}^{D}\sum_{i=1}^{D}\sum_{\varphi\in\mathcal{D}^{|\alpha|}}\frac{\sigma(\varphi)!}{\alpha!}\Pi_g(\varsigma(\alpha),\varphi)\cdot(\sigma_i(\varphi)+1)\frac{\partial u_d}{\partial x_i}f_{\sigma(\varphi)+e_i-e_d}.
$$

(4.57)

This proves the first equality. The second equality can be similarly proved, and the details are omitted. $\square$

Now the proof of Theorem 3 is given as follows:

Proof of Theorem 3. Since $(n_1,\cdots,n_D)$ is a unit vector, we let $\mathbf{G}=(g_{ij})_{D\times D}$ be an orthogonal matrix with its first row as $(n_1,\cdots,n_D)$. Now we use this matrix as the rotation matrix and define $\tilde{\boldsymbol{w}}$ as (4.36) and (4.32). It is obvious that the relation between $\tilde{\boldsymbol{w}}$ and $\boldsymbol{w}$ is linear. Therefore, there exists a constant matrix $\mathbf{R}$ (see (4.36)) depending on $\mathbf{G}$ such that

$$
\tilde{\boldsymbol{w}} = \mathbf{R}\boldsymbol{w}, \tag{4.58}
$$

and $\mathbf{R}$ is invertible since $\boldsymbol{w}$ can be obtained from $\tilde{\boldsymbol{w}}$ by applying the rotation matrix $\mathbf{G}^{-1}$. Lemma 7–9 have clearly shown that the "rotated equations"

$$
\mathbf{T}(\tilde{\boldsymbol{w}})\frac{\partial\tilde{\boldsymbol{w}}}{\partial t} + \sum_{j=1}^{D}\mathbf{T}(\tilde{\boldsymbol{w}})\hat{\mathbf{M}}_j(\tilde{\boldsymbol{w}})\frac{\partial\tilde{\boldsymbol{w}}}{\partial\tilde{x}_j} = 0 \tag{4.59}
$$

can be deduced from (4.6) by linear operations. Thus there exists a square matrix $\mathbf{H}(\boldsymbol{w})$ such that

$$\mathbf{H}(\boldsymbol{w})\mathbf{T}(\boldsymbol{w})\frac{\partial \boldsymbol{w}}{\partial t} + \sum_{j=1}^{D} \mathbf{H}(\boldsymbol{w})\mathbf{T}(\boldsymbol{w})\hat{\mathbf{M}}_j(\boldsymbol{w})\frac{\partial \boldsymbol{w}}{\partial x_j} = 0 \tag{4.60}$$

is identical to (4.59). Matching the terms with time derivatives, one finds $\mathbf{H}(\boldsymbol{w}) = \mathbf{T}(\tilde{\boldsymbol{w}})\mathbf{R}\mathbf{T}(\boldsymbol{w})^{-1}$. Thus (4.60) becomes

$$\mathbf{T}(\tilde{\boldsymbol{w}})\frac{\partial \tilde{\boldsymbol{w}}}{\partial t} + \sum_{j=1}^{D} \mathbf{T}(\tilde{\boldsymbol{w}})\mathbf{R}\hat{\mathbf{M}}_j(\boldsymbol{w})\frac{\partial \boldsymbol{w}}{\partial x_j} = 0. \tag{4.61}$$

Using (4.40), the above equation can be rewritten as

$$\mathbf{T}(\tilde{\boldsymbol{w}})\mathbf{R}\frac{\partial \boldsymbol{w}}{\partial t} + \sum_{j=1}^{D}\sum_{d=1}^{D} g_{dj}\mathbf{T}(\tilde{\boldsymbol{w}})\mathbf{R}\hat{\mathbf{M}}_j(\boldsymbol{w})\frac{\partial \boldsymbol{w}}{\partial \tilde{x}_d} = 0. \tag{4.62}$$

Compared with (4.59), one concludes

$$\sum_{j=1}^{D} g_{1j}\mathbf{T}(\tilde{\boldsymbol{w}})\mathbf{R}\hat{\mathbf{M}}_j(\boldsymbol{w}) = \mathbf{T}(\tilde{\boldsymbol{w}})\hat{\mathbf{M}}_1(\mathbf{R}\boldsymbol{w})\mathbf{R}. \tag{4.63}$$

Multiplying both sides by $\mathbf{R}^{-1}\mathbf{T}(\tilde{\boldsymbol{w}})^{-1}$, (4.7) is attained. Recalling $\hat{\mathbf{M}}_1 = \hat{\mathbf{A}}_M$ and that the first component of the macroscopic velocity after the rotation is $\boldsymbol{u} \cdot \boldsymbol{n}$ (see (4.32)), the diagonalizability and the eigenvalues of the matrix (4.7) are naturally obtained using Theorem 2. $\qquad\square$

5 Riemann Problem

Though the regularized moment system (4.3) is given by the moment expansion up to an arbitrary order M thus extremely complex, we can clarify appreciably the structures of the elementary waves of this system with Riemann initial value, including the rarefaction wave, contact discontinuity and shock wave. Definitely, the structure of the elementary wave is fundamental for further investigation into the behavior of the solution of the system. Furthermore, the solution structure of the Riemann problem is instructional for studying the approximate Riemann solver, which is the basis of the numerical methods using Godunov type schemes. The analysis below shows that the structure of the elementary wave of the Riemann problem is quite natural an extension of that of Euler equations, which indicates that the regularized moment system (4.3) is actually a very reasonable high order moment approximation of Boltzmann equation. Following [12] where the multi-dimensional Euler equations are studied, we consider the x_1-split, D-dimensional Riemann problem as below:

$$\begin{cases} \dfrac{\partial \boldsymbol{w}}{\partial t} + \hat{\mathbf{A}}_M\dfrac{\partial \boldsymbol{w}}{\partial x_1} = 0, \\[2mm] \boldsymbol{w}(x_1, t=0) = \begin{cases} \boldsymbol{w}_L & \text{if } x_1 < 0, \\ \boldsymbol{w}_R & \text{if } x_1 > 0. \end{cases} \end{cases} \tag{5.1}$$

The Riemann problem with 1D velocity space has been studied in [1] in detail. Here we focus on the case of $D \geq 2$.

Let us first recall the definition of the notations $\tilde{\mathbf{A}}_M$, $\hat{\mathbf{A}}_M$, $\mathbf{B}$, $\hat{\mathbf{B}}$, $\boldsymbol{w}$, $\boldsymbol{r}_{\hat{\alpha},i}$, $\hat{\boldsymbol{r}}_{\hat{\alpha},i}$, $\boldsymbol{v}^{(j)}$ and $\lambda_{i,k}$ in Section 3. In particular, we need the expressions of $\hat{\mathbf{B}}\mathbf{B}^{-1}$ and $\boldsymbol{v}^{(j)}$, which read

$$
\hat{\mathbf{B}}\mathbf{B}^{-1} = \begin{bmatrix} \mathbf{I} & 0 & 0 \\ -\mathbf{B}_{21} & \mathbf{I} & 0 \\ * & * & \mathbf{I} \end{bmatrix} \tag{3.58'}
$$

and for any $|\alpha| = M$, $k = \alpha_1 + 1$, $j = \mathcal{N}_{D-1}(\hat{\alpha})$,

$$
\boldsymbol{v}^{(j)} = \hat{\mathbf{B}}\mathbf{B}^{-1}I_j, \quad He_k(\mathrm{C}) = 0, \tag{3.59'}
$$

where I_j is the j-th column of the $N_v \times N_v$ identity matrix. $\hat{\boldsymbol{r}}_{\hat{\alpha},i}$, which depends on $\boldsymbol{v}^{(j)}$ and $\lambda_{i,k} = u_1 + \mathrm{C}_{i,k}\sqrt{\theta}$, is the eigenvector of $\hat{\mathbf{A}}_M$ for the eigenvalue $\lambda_{i,k} = u_1 + \mathrm{C}_{i,k}\sqrt{\theta}$, where $j = \mathcal{N}_{D-1}(\hat{\alpha})$, $k = M + 1 - |\hat{\alpha}|$. As the first conclusion on the Riemann problem (5.1), we have the following theorem:

Theorem 4. *Each characteristic field of* (5.1) *is either genuinely nonlinear or linearly degenerate. And one characteristic field is genuinely nonlinear if and only if $\boldsymbol{v}$ (determined by the right eigenvector through* (3.23)*) and the eigenvalue $\lambda = u_1 + \mathrm{C}\sqrt{\theta}$ satisfy one of the following two conditions:*

1. *$\boldsymbol{v} = \boldsymbol{v}^{(1)}$, and C subject to $He_{M+1}(\mathrm{C}) = 0$ and $\mathrm{C} \neq 0$;*

2. *$\boldsymbol{v} = \boldsymbol{v}^{(j)}$, $j = \mathcal{N}_{D-1}(2\hat{e}_k)$, $k \in \mathcal{D}\backslash\{1\}$, and C subject to $He_{M-1}(\mathrm{C}) = 0$ and $\mathrm{C} \neq 0$.*

Proof. Let $\hat{\boldsymbol{r}}$ denote an eigenvector of $\hat{\mathbf{A}}_M$ with the eigenvalue $\lambda = u_1 + \mathrm{C}\sqrt{\theta}$ and $\boldsymbol{v}$ is the corresponding vector determined by (3.23). Since

$$
\lambda = u_1 + \mathrm{C}\sqrt{\theta} = u_1 + \mathrm{C}\sqrt{\frac{\sum_{d=1}^{D} p_{2e_d}}{D\rho}} \tag{5.2}
$$

depends only on ρ, u_1 and $p_{2e_d}/2$, $d \in \mathcal{D}$, we have

$$
\begin{aligned}
\nabla_{\boldsymbol{w}}\lambda \cdot \hat{\boldsymbol{r}} &= -\frac{\mathrm{C}\sqrt{\theta}}{2\rho} \cdot \rho r_\rho + 1 \cdot \mathrm{C}\sqrt{\theta}r_\rho + \frac{\mathrm{C}}{D\rho\sqrt{\theta}} \cdot \frac{\mathrm{C}^2\theta}{2}\rho r_\rho + \sum_{d=2}^{D} \frac{\mathrm{C}}{D\rho\sqrt{\theta}} \cdot \rho\theta r_{p_{2e_d}/2} \\
&= \frac{\sqrt{\theta}\mathrm{C}}{2}\left[\left(1 + \frac{\mathrm{C}^2}{D}\right)v_1 + \sum_{d=2}^{D} \frac{2}{D}v_{\mathcal{N}_{D-1}(2\hat{e}_d)}\right].
\end{aligned} \tag{5.3}
$$

- If $\boldsymbol{v} = \boldsymbol{v}^{(1)}$, then (3.59') shows $\boldsymbol{v}^{(1)} = \hat{\mathbf{B}}\mathbf{B}^{-1}I_1$ and $He_{M+1}(\mathrm{C}) = 0$. From (3.58'), we get

$$
v_1 = 1 \text{ and } v_{\mathcal{N}_{D-1}(2\hat{e}_2)} = \cdots = v_{\mathcal{N}_{D-1}(2\hat{e}_D)} = \frac{\mathrm{C}^2}{2D} - \frac{1}{2}. \tag{5.4}
$$

Thus (5.3) can be written as

$$
\nabla_{\boldsymbol{w}}\lambda \cdot \hat{\boldsymbol{r}} = \frac{(D+1)\sqrt{\theta}\mathrm{C}}{4D^2}\left(\mathrm{C}^2 + D\right). \tag{5.5}
$$

Hence,

$$
\begin{cases} \nabla_{\boldsymbol{w}}\lambda \cdot \hat{\boldsymbol{r}} \equiv 0, & \text{if } \mathrm{C} = 0, \\ \nabla_{\boldsymbol{w}}\lambda \cdot \hat{\boldsymbol{r}} \not\equiv 0, & \text{otherwise.} \end{cases}
$$

- If $\boldsymbol{v} = \boldsymbol{v}^{(j)}$, $j = \mathcal{N}_{D-1}(2\hat{e}_k)$ for any $k \in \mathcal{D}\backslash\{1\}$, (3.59') shows $\boldsymbol{v} = \hat{\mathbf{B}}\mathbf{B}^{-1}I_j$ and $He_{M-1}(\mathrm{C}) = 0$. From (3.58'), we can get

$$v_j = 1 \text{ and } v_1 = v_l = 0, \ l = \mathcal{N}_{D-1}(2\hat{e}_d) \text{ for any } d \in \mathcal{D}\backslash\{1, k\}.$$

  Then (5.3) is simplified as

$$\nabla_{\boldsymbol{w}}\lambda \cdot \hat{\boldsymbol{r}}_{\hat{\alpha},i} = \frac{\sqrt{\theta}}{D}\mathrm{C}. \tag{5.6}$$

  Again, we have

$$\begin{cases} \nabla_{\boldsymbol{w}}\lambda \cdot \hat{\boldsymbol{r}} \equiv 0, & \text{if } \mathrm{C} = 0, \\ \nabla_{\boldsymbol{w}}\lambda \cdot \hat{\boldsymbol{r}} \not\equiv 0, & \text{otherwise.} \end{cases}$$

- Otherwise, (3.58') indicates $v_1 = v_{\mathcal{N}_{D-1}(2\hat{e}_k)} = 0$ for each $k \in \mathcal{D}\backslash\{1\}$. Hence $\nabla_{\boldsymbol{w}}\lambda\cdot\hat{\boldsymbol{r}} \equiv 0$ always holds.

This completes the proof. $\qquad\square$

This theorem reveals that for each characteristic field, the eigenvalue is constant or varies monotonically along the integral curve, resulting in simple wave structures. Below, some elementary waves including the rarefaction waves, contact discontinuities and shock waves are studied in detail, and the basic relations across these waves are established.

The analysis below is based on the fact that an eigenvector $\hat{\boldsymbol{r}}$ of $\hat{\mathbf{A}}_M$ for the eigenvalue $\lambda = u_1 + \mathrm{C}\sqrt{\theta}$ depends only on $\boldsymbol{v}$ and C. With Theorem 4 and the forms of $\mathbf{B}$ and $\hat{\mathbf{B}}\mathbf{B}$ in Lemma 2, we can divide characteristic fields into three cases:

Case 1: $\boldsymbol{v} = \boldsymbol{v}^{(1)}$, and C subject to $He_{M+1}(\mathrm{C}) = 0$, and $\mathrm{C} \neq 0$.

Case 2: $\boldsymbol{v} = \boldsymbol{v}^{(j)}$, $j = \mathcal{N}_{D-1}(2\hat{e}_k)$ for any $k \in \mathcal{D}\backslash\{1\}$, and C subject to $He_{M-1}(\mathrm{C}) = 0$, and $\mathrm{C} \neq 0$.

Case 3: otherwise.

For convenience, let characteristic field α denote the characteristic field corresponding to the eigenvector $\hat{\boldsymbol{r}}_{\hat{\alpha},i}$ for the eigenvalue $\lambda_{i,k} = u_1 + \mathrm{C}_{i,k}\sqrt{\theta}$ with $i = \alpha_1$, $k = M + 1 - |\hat{\alpha}|$. Below, the rarefaction waves, contact discontinuities and shock waves will be studied respectively.

5.1 Rarefaction waves

For the regularized moment system, if two states $\boldsymbol{w}^L$ and $\boldsymbol{w}^R$ are connected by a rarefaction wave in a genuinely nonlinear field α, then the following two conditions must be met:

- constancy of the *generalised Riemann invariants* across the wave, saying the integral curve $\tilde{\boldsymbol{w}}(\zeta) = (\tilde{w}_1(\zeta), \tilde{w}_2(\zeta), \cdots, \tilde{w}_N(\zeta))$ in the N-dimensional phase space satisfies

$$\tilde{\boldsymbol{w}}'(\zeta) = \hat{\boldsymbol{r}}_{\hat{\alpha},i}(\tilde{\boldsymbol{w}}), \tag{5.7}$$

  with $i = \alpha_1$.

- divergence of characteristics

$$\lambda_{i,k}(\boldsymbol{w}^L) < \lambda_{i,k}(\boldsymbol{w}^R). \tag{5.8}$$

Fortunately, for a given point $\boldsymbol{w}^0 = (\rho^0, u_1^0, \cdots, w_j^0, \cdots, w_N^0)$ in the phase space, the integral curve across $\boldsymbol{w}^0$ can be given. Since $p = \dfrac{1}{D}\displaystyle\sum_{d=1}^{D} p_{2e_d}$, we let $p^0 = \dfrac{1}{D}\displaystyle\sum_{d=1}^{D} p_{2e_d}^0$. The results are rather tedious, and here the integral curves are only partially given in three cases as below:

- If $\boldsymbol{v} = \boldsymbol{v}^{(1)}$, we have

$$r_\rho = \rho, \quad r_{u_1} = \mathrm{C}_{i,k}\sqrt{\theta}, \quad r_{u_d} = 0,$$

$$r_{p_{2e_1}/2} = \frac{\mathrm{C}_{i,k}^2}{2}\rho\theta, \quad r_{p_{2e_d}/2} = \frac{\mathrm{C}_{i,k}^2 - D}{2D}\rho\theta, \quad d \in \mathcal{D}\backslash\{1\}.$$

Let $\Gamma = \dfrac{D - 1 + \mathrm{C}_{i,k}^2}{2D - 1}$, and then we have

$$\tilde{\rho}(\zeta) = \rho^0 \exp(\zeta), \tag{5.9a}$$

$$\tilde{u}_1(\zeta) = u_1^0 + \frac{2\mathrm{C}_{i,k}\sqrt{\theta^0}}{\Gamma - 1}\left[\exp\left(\frac{\Gamma - 1}{2}\zeta\right) - 1\right], \tag{5.9b}$$

$$\tilde{u}_d(\zeta) = u_d^0, \quad d = 2, \cdots, D, \tag{5.9c}$$

$$\tilde{p}_{2e_1}(\zeta) = p_{2e_1}^0 + \frac{\mathrm{C}_{i,k}^2 - D}{D\Gamma}p^0\left[\exp\left(\Gamma\zeta\right) - 1\right], \tag{5.9d}$$

$$\tilde{p}(\zeta) = p^0 \exp\left(\Gamma\zeta\right). \tag{5.9e}$$

- If $\boldsymbol{v} = \boldsymbol{v}^{(j)}$, $j = \mathcal{N}_{D-1}(2\hat{e}_k)$, $k \in \mathcal{D}\backslash\{1\}$, we have

$$r_\rho = 0, \quad , r_{u_d} = 0, \ d \in \mathcal{D}, \quad r_{p_{2e_k}/2} = \rho\theta, \quad r_{p_{2e_d}/2} = 0, \ d \in \mathcal{D}\backslash\{k\}.$$

Hence, the integral curve satisfies

$$\tilde{\rho}(\zeta) = \rho^0, \tag{5.10a}$$

$$\tilde{u}_d(\zeta) = u_d^0, \quad d = 1, \cdots, D, \tag{5.10b}$$

$$\tilde{p}_{2e_1}(\zeta) = p_{2e_1}^0, \tag{5.10c}$$

$$\tilde{p}(\zeta) = p^0 \exp\left(\frac{2\zeta}{D}\right). \tag{5.10d}$$

- Otherwise ($\boldsymbol{v} = \boldsymbol{v}^{(j)}$, $j \neq \mathcal{N}_{D-1}(2\hat{e}_k)$ for any $k \in \mathcal{D}$),

$$r_\rho = r_{u_1} = r_{p_{2e_d}/2} = 0, \quad d \in \mathcal{D}.$$

Hence, we have

$$\tilde{\rho}(\zeta) = \rho^0, \qquad\qquad \tilde{u}_1(\zeta) = u_1^0, \tag{5.11a}$$

$$\tilde{p}_{2e_1}(\zeta) = p_{2e_1}^0, \qquad\qquad \tilde{p}(\zeta) = p^0. \tag{5.11b}$$

One can check that (5.9), (5.11) and (5.10) satisfy (5.7). And an eigenvalue, which satisfies (3.69), of $\hat{\mathbf{A}}_M(\tilde{\boldsymbol{w}}(\zeta))$ is as

$$s_{i,k}(\tilde{\boldsymbol{w}}(\zeta)) = \tilde{u}_1(\zeta) + \mathrm{C}_{i,k}\sqrt{\tilde{p}(\zeta)/\tilde{\rho}(\zeta)}$$

$$= \begin{cases} u_1^0 + \mathrm{C}_{i,k}\sqrt{\theta^0} + \dfrac{\Gamma+1}{\Gamma-1}\mathrm{C}_{i,k}\sqrt{\theta^0}\left[\exp\left(\dfrac{\Gamma-1}{2}\zeta\right)-1\right], & \text{for } \boldsymbol{v}^{(1)}, \\[2ex] u_1^0 + \mathrm{C}_{i,k}\sqrt{\theta^0}\exp\left(\dfrac{\zeta}{D}\right), & \text{for } \boldsymbol{v}^{(j)},\ j = \mathcal{N}_{D-1}(2\hat{e}_k),\ k \in \mathcal{D}\backslash\{1\}, \\[2ex] u_1^0 + \mathrm{C}_{i,k}\sqrt{\theta^0}, & \text{otherwise.} \end{cases}$$

It is convenient to verify that $s_{i,k}(\tilde{\boldsymbol{w}}(\zeta)) \gtrless s_{i,k}(\boldsymbol{w}^0)$ if and only if $\mathrm{C}_{i,k}\zeta \gtrless 0$, and $\boldsymbol{v}$ and $\mathrm{C}_{i,k}$ satisfy case 1 or case 2. Therefore, if the left state $\boldsymbol{w}^L$ and the right state $\boldsymbol{w}^R$ are connected by a rarefaction wave and let $\boldsymbol{w}^0 = \boldsymbol{w}^L$, (5.8) indicates $s_{i,k}(\boldsymbol{w}^L) < s_{i,k}(\boldsymbol{w}^R)$, hence $\mathrm{C}_{i,k}\zeta > 0$ and $\boldsymbol{v}, \mathrm{C}_{i,k}$ satisfies case 1 or case 2. Therefore, we have that

- for case 1:
$$u_d^L = u_d^R, \quad d = 2,\cdots,D,$$
and

$$\text{if } \mathrm{C}_{i,k} > 0, \text{ then } u_1^L < u_1^R, \quad p^L < p^R, \tag{5.12a}$$
$$\text{if } \mathrm{C}_{i,k} < 0, \text{ then } u_1^L < u_1^R, \quad p^L > p^R. \tag{5.12b}$$

- for case 2:
$$u_d^L = u_d^R, \quad d = 2,\cdots,D,$$
and

$$\text{if } \mathrm{C}_{i,k} > 0, \text{ then } u_1^L = u_1^R, \quad p^L < p^R, \tag{5.13a}$$
$$\text{if } \mathrm{C}_{i,k} < 0, \text{ then } u_1^L = u_1^R, \quad p^L > p^R. \tag{5.13b}$$

5.2 Contact discontinuities

For a contact discontinuities, (5.7) is still valid, and the divergence of characteristics is replaced by

$$\lambda_{i,k}(\boldsymbol{w}^L) = \lambda_{i,k}(\boldsymbol{w}^R). \tag{5.14}$$

According to Theorem 4 and analysis in Section 5.1, the contact discontinuities can be founded if and only if $\boldsymbol{v}$ and $\mathrm{C}_{i,k}$ satisfy case 3.

- For $\boldsymbol{v}^{(1)}$, (5.14) means $\mathrm{C}_{i,k} = 0$. Substituting it into (5.9), we can get u_d, $d \in \mathcal{D}$ are invariant, while p, p_{2e_1} are not (otherwise, (5.9e) gives us $\zeta = 0$, thus $\boldsymbol{w}^L = \boldsymbol{w}^R$).

- For $\boldsymbol{v}^{(j)}$, $j = \mathcal{N}_{D-1}(2\hat{e}_k)$, $k \in \mathcal{D}\backslash\{1\}$, (5.14) means $\mathrm{C}_{i,k} = 0$ again. (5.10) shows ρ, u_d, $d \in \mathcal{D}$, p_{2e_1} are invariant, while p is not (otherwise, (5.10d) gives us $\zeta = 0$, which results $\boldsymbol{w}^L = \boldsymbol{w}^R$).

- Otherwise, (5.11) shows u_1, p, p_{2e_1} are all invariant.

Summarizing the discussion above, we conclude that if $\mathrm{C}_{i,k} \neq 0$, then u_1, p, p_{2e_1} are invariant across the contact discontinuities, while if $\mathrm{C}_{i,k} = 0$, u_1 is invariant and p is not. However, u_d, $d = 2,\cdots,D$ may change discontinuously across a contact discontinuity. In fact, the case $\boldsymbol{v} = \mathbf{B}^{-1}I_d$, $d = 2,\cdots,D$ corresponds to a contact discontinuity where u_d is discontinuous. This is similar as the Euler equations.

5.3 Shock waves

The discussion of the shock wave needs some more scrupulosity. As is well known, the jump condition on the shock wave is sensitive to the form of the hyperbolic equations. Thus, before we give the Rankine-Hugoniot condition, it is necessary to rewrite (5.1) in an appropriate form. However, (5.1) cannot be written as conservation laws due to the presence of $\mathcal{R}_{M,D}^1(\alpha)$. Nevertheless, (5.1) can still keep the conservation of the conservative moments with orders from 0 to $M-1$. Therefore, (5.1) can be reformulated by $\mathcal{N}_D((M-1)e_D)$ conservation laws and $N - \mathcal{N}_D((M-1)e_D)$ non-conservative equations.

Let

$$\boldsymbol{F} = (F_0, F_{e_1}, F_{e_2}, \cdots, F_{Me_D}), \quad F_\alpha = \frac{1}{\alpha!} \int_{\mathbb{R}^D} \boldsymbol{\xi}^\alpha f \, \mathrm{d}\boldsymbol{\xi}, \quad |\alpha| \leq M, \qquad (5.15)$$

where $\boldsymbol{\xi}^\alpha = \prod_{d=1}^D \xi_d^{\alpha_d}$ and F_0 stands for $F_\alpha|_{\alpha=\boldsymbol{0}}$. Then (5.1) can be written as

$$\begin{aligned}
\frac{\partial F_\alpha}{\partial t} + (\alpha_1 + 1)\frac{\partial F_{\alpha+e_1}}{\partial x_1} &= 0, \quad |\alpha| < M, \\
\frac{\partial F_\alpha}{\partial t} + (\alpha_1 + 1)\frac{\partial \hat{F}_\alpha}{\partial x_1} - \mathcal{R}_{M,D}^1(\alpha) &= 0, \quad |\alpha| = M.
\end{aligned} \qquad (5.16)$$

The relation between $\boldsymbol{F}$ and $\boldsymbol{w}$ is

$$f_\alpha = \sum_{|\beta| \leq |\alpha|} (-1)^{|\alpha-\beta|} F_\beta \frac{He_{\alpha-\beta}\left(\dfrac{\boldsymbol{u}}{\sqrt{\theta}}\right)}{(\alpha-\beta)!} \theta^{|\alpha-\beta|/2}, \qquad (5.17)$$

$$u_i = \frac{F_{e_i}}{F_0}, \quad p_{2e_i} = 2F_{2e_i} - \frac{F_{e_i}^2}{F_0}, \quad i \in \mathcal{D},$$

where $He_\alpha\left(\dfrac{\boldsymbol{u}}{\sqrt{\theta}}\right) = \prod_{d=1}^D He_{\alpha_d}\left(\dfrac{u_d}{\sqrt{\theta}}\right)$ and $He_\alpha(\boldsymbol{x}) = 0$ if at least one α_j is negative. In addition, $\hat{F}_\alpha$ is as

$$\hat{F}_\alpha = \sum_{|\beta| \leq |\alpha|} (-1)^{|\alpha-\beta|} F_\beta \frac{He_{\alpha+e_1-\beta}\left(\dfrac{\boldsymbol{u}}{\sqrt{\theta}}\right)}{(\alpha+e_1-\beta)!} \theta^{|\alpha+e_1-\beta|/2}. \qquad (5.18)$$

For convenience, the quasi-linear form of (5.16) is written as

$$\frac{\partial \boldsymbol{F}}{\partial t} + \boldsymbol{\Gamma}(\boldsymbol{F})\frac{\partial \boldsymbol{F}}{\partial x_1} = 0, \qquad (5.19)$$

where $\boldsymbol{\Gamma}(\boldsymbol{F})$ is an $N \times N$ matrix and depends on (5.16).

Since (5.19) is not a conservative system, we have to adopt the DLM theory [10] to study the shock wave. For a shock wave the two constant states $\boldsymbol{F}^L$ and $\boldsymbol{F}^R$ are connected through a single jump discontinuity in a genuinely non-linear field α travelling at the speed S_α, and the following two conditions apply

- Generalized Rankine-Hugoniot condition:

$$\int_0^1 \left[S_\alpha \boldsymbol{I} - \boldsymbol{\Gamma}\left(\boldsymbol{\Phi}(\nu; \boldsymbol{F}^L, \boldsymbol{F}^R)\right)\right] \frac{\partial \boldsymbol{\Phi}}{\partial \nu}(\nu; \boldsymbol{F}^L, \boldsymbol{F}^R) \, \mathrm{d}\nu = 0, \qquad (5.20)$$

where $\mathbf{I}$ is the $N \times N$ identity matrix, and $\boldsymbol{\Phi}(\nu; \boldsymbol{F}^L, \boldsymbol{F}^R)$ is a locally Lipschitz mapping satisfying

$$\boldsymbol{\Phi}(0; \boldsymbol{F}^L, \boldsymbol{F}^R) = \boldsymbol{F}^L, \quad \boldsymbol{\Phi}(1; \boldsymbol{F}^L, \boldsymbol{F}^R) = \boldsymbol{F}^R. \tag{5.21}$$

We refer the readers to [10] for details.

- Entropy condition:

$$\lambda_{i,k}(\boldsymbol{F}^L) > S_\alpha > \lambda_{i,k}(\boldsymbol{F}^R), \tag{5.22}$$

where $i = \alpha_1$ and $k = M + 1 - |\hat{\alpha}|$.

For conservation laws, (5.20) is the same as the classical Rankine-Hugoniot condition. Thus the first $\mathcal{N}_D((M-1)e_D)$ rows of (5.20) are independent of $\boldsymbol{\Phi}$. This allows us to analyze the properties of the shock waves without regarding the form of $\boldsymbol{\Phi}$.

The first equation and the $(D+1)$-th equation of (5.20) are precisely as

$$\rho^L u_1^L - \rho^R u_1^R = S_\alpha(\rho^L - \rho^R), \tag{5.23}$$

$$\rho^L(u_1^L)^2 + p_{2e_1}^L - \rho^R(u_1^R)^2 - p_{2e_1}^R = S_\alpha(\rho^L u_1^L - \rho^R u_1^R). \tag{5.24}$$

- If $\rho^L \neq \rho^R$, (5.23) and (5.24) give

$$S_\alpha = \frac{\rho^L u_1^L - \rho^R u_1^R}{\rho^L - \rho^R} \tag{5.25a}$$

$$= \frac{\rho^L(u_1^L)^2 + p_{2e_1}^L - \rho^R(u_1^R)^2 - p_{2e_1}^R}{\rho^L u_1^L - \rho^R u_1^R}. \tag{5.25b}$$

Substituting (5.25a) into (5.22) and multiplying both sides with $(\rho^L - \rho^R)^2$, we get

$$\rho^L(u_1^L - u_1^R)(\rho^L - \rho^R) > \mathrm{C}_{i,k}(\rho^L - \rho^R)^2\sqrt{\theta^R}, \tag{5.26a}$$

$$\rho^R(u_1^L - u_1^R)(\rho^L - \rho^R) < \mathrm{C}_{i,k}(\rho^L - \rho^R)^2\sqrt{\theta^L}. \tag{5.26b}$$

If $\mathrm{C}_{i,k} > 0$, (5.26a) gives

$$(u_1^L - u_1^R)(\rho^L - \rho^R) > 0. \tag{5.27}$$

Therefore, we can divide (5.26) by $(u_1^L - u_1^R)(\rho^L - \rho^R)$ and obtain

$$\frac{\rho^L}{\sqrt{\theta^R}} > \frac{\mathrm{C}_{i,k}(\rho^L - \rho^R)}{u_1^L - u_1^R} > \frac{\rho^R}{\sqrt{\theta^L}}. \tag{5.28}$$

Thus we have

$$(\rho^L)^2\theta^L > (\rho^R)^2\theta^R. \tag{5.29}$$

Furthermore, (5.25) gives us the relation

$$(\rho^L - \rho^R)(p_{2e_1}^L - p_{2e_1}^R) = \rho^L\rho^R(u_1^L - u_1^R)^2. \tag{5.30}$$

If $\rho^L < \rho^R$, (5.27) indicates $u_1^L < u_1^R$. (5.29) can be written as $\rho^L p^L > \rho^R p^R$, so we have $p^L > p^R$. If $\rho^L > \rho^R$, (5.27) indicates $u_1^L > u_1^R$. Summarizing these results, we get

$$\begin{aligned} &\text{if } \rho^L < \rho^R, \text{ then } u_1^L < u_1^R \text{ and } p^L > p^R, \\ &\text{if } \rho^L > \rho^R, \text{ then } u_1^L > u_1^R. \end{aligned} \tag{5.31}$$

36

Analogously, if $C_{i,k} < 0$, we have

$$(u_1^L - u_1^R)(\rho^L - \rho^R) < 0, \quad (\rho^L)^2 \theta^L < (\rho^R)^2 \theta^R, \tag{5.32}$$

and (5.30) still holds. Hence we get that

$$\begin{aligned} &\text{if } \rho^L > \rho^R, \text{ then } u_1^L < u_1^R \text{ and } p^L < p^R, \\ &\text{if } \rho^L < \rho^R, \text{ then } u_1^L > u_1^R. \end{aligned} \tag{5.33}$$

- If $\rho^L = \rho^R$, (5.23) and (5.24) make that $u_1^L = u_1^R$ and $p_{2e_1}^L = p_{2e_1}^R$, respectively. Therefore, (5.22) is turned into

$$C_{i,k}\sqrt{\theta^L} > C_{i,k}\sqrt{\theta^R}. \tag{5.34}$$

The following result is then attained

$$\text{if } C_{i,k} > 0, \text{ then } u_1^L = u_1^R, \quad p^L > p^R, \tag{5.35a}$$

$$\text{if } C_{i,k} < 0, \text{ then } u_1^L = u_1^R, \quad p^L < p^R. \tag{5.35b}$$

Now we summarize all our discussion on the entropy conditions of the three types of elementary waves in the following theorem:

Theorem 5. *For the Riemann problem* (5.1), *for the wave of the α-th family, $C_{i,k}$, the macroscopic velocities and pressures on both sides of the wave have the relation with the type of the wave as in Table 1, where $C_{i,k}$ corresponds to the eigenvalue $\lambda_{i,k} = u_1 + C_{i,k}\sqrt{\theta}$, $i = \alpha_1$ and $k = M + 1 - |\hat{\alpha}|$.*

Wave type	Eigenvalue	Velocity and Pressure
Rarefaction wave	$C_{i,k} > 0$	$u_1^L \leq u_1^R, p^L < p^R$
	$C_{i,k} < 0$	$u_1^L \leq u_1^R, p^L > p^R$
Shock wave	$C_{i,k} > 0$	$u_1^L \leq u_1^R, p^L > p^R$
	$C_{i,k} < 0$	$u_1^L \leq u_1^R, p^L < p^R$
	$C_{i,k} \neq 0$	$u_1^L > u_1^R$
Contact discontinuity	$C_{i,k} = 0$	$u_1^L = u_1^R$
	$C_{i,k} \neq 0$	$u_1^L = u_1^R, p^L = p^R$

Table 1: The relation between the type classification of elementary wave and the eigenvalue, macroscopic velocity and pressure.

Acknowledgements

Appendix

A   Collection of Notations

We list below some of the notations used for convenience.

$Q(f,f)$ The Boltzmann collision operator taken as zero in this paper.

$\mathrm{C}_{i,k}$ The i-th zero of $He_k(x)$.

α D-dimensional multi-index.

$\tilde{\alpha}$ D-dimensional multi-index, defined as $\tilde{\alpha} = \alpha - \alpha_1 e_1$.

$\hat{\alpha}$ $(D-1)$-dimensional multi-index, defined by (3.6b).

$\mathcal{S}_{D,M}$ All the D-dimensional multi-index satisfying the sum of components not more than M, defined by (3.6c).

$\mathcal{S}_{D,M}(\hat{\alpha})$ The subset of $\mathcal{S}_{D,M}$, defined by (3.6c).

$\mathcal{N}_D(\alpha)$ The ordinal number of α in $\mathcal{S}_{D,M}$, which is permuted by lexicographic order, defined by (3.7).

N Cardinality of $\mathcal{S}_{D,M}$.

$\boldsymbol{w}$ The basic moments in the moment system, defined by (3.9).

$\mathbf{A}_M$ The coefficient matrix of the moment system, defined by (3.10).

$\mathbf{I}$ Identity matrix.

$\mathcal{R}_{M,D}$ Regularization term based on the characteristic speed correction.

$\mathcal{R}_{M,D}^j$ Regularization term in x_j direction.

$\hat{\mathbf{A}}_M$ The regularized matrix of $\mathbf{A}_M$, defined by (3.19).

$\boldsymbol{\Lambda}$ The diagonal matrix, used to make $\hat{\mathbf{A}}_M$ dimensionless, defined by (3.21).

$\tilde{\mathbf{A}}_M$ The dimensionless matrix of $\hat{\mathbf{A}}_M - u_1\mathbf{I}$.

$\hat{p}_{e_i+e_k}, g_\alpha$ The dimensionless of $p_{e_i+e_k}$, f_α.

$\boldsymbol{r}$ The right eigenvector of $\tilde{\mathbf{A}}_M$, defined by (3.30).

$\boldsymbol{v}$ The components of it is a subset of that of $\boldsymbol{r}$, and $v_{\mathcal{N}_{D-1}(\hat{\alpha})} = r_{w_{\mathcal{N}_D(\tilde{\alpha})}}$.

$\mathbf{B}$ The coefficient matrix defined in Lemma 2.

$\hat{\mathbf{B}}$ The block diagonal matrix, used to make $\mathbf{B}$ a unit lower triangular matrix, defined in Lemma 2.

$\boldsymbol{v}^{(j)}$ The basis of $\boldsymbol{v}$, defined by (3.59).

$\boldsymbol{r}_{\hat{\alpha},i}$ The eigenvector of $\tilde{\mathbf{A}}_M$, depended on $\boldsymbol{v}^{(j)}$, $j = \mathcal{N}_{D-1}(\hat{\alpha})$, for the eigenvalue $\mathrm{C}_{i,k}$, $k = M + 1 - |\hat{\alpha}|$.

$\tilde{\mathcal{P}}_{D,M}$ The characteristic polynomial of $\tilde{\mathbf{A}}_M$, defined by (3.63).

$\lambda_{i,k}$ The eigenvalue of $\hat{\mathbf{A}}_M$, $\lambda_{i,k} = u_1 + \mathrm{C}_{i,k}\sqrt{\theta}$.

$\hat{\boldsymbol{r}}_{\hat{\alpha},i}$ The eigenvector of $\hat{\mathbf{A}}_M$, corresponding to $\boldsymbol{r}_{\hat{\alpha},i}$, for the eigenvalue $\mathrm{C}_{i,k}$, $k = M + 1 - |\hat{\alpha}|$, defined by (3.69).

$\mathcal{P}_{D,M}$ The characteristic polynomial of $\hat{\mathbf{A}}_M$, defined by (3.67).

$\mathbf{G} = (g_{ij})_{D \times D}$	The rotational matrix.
$\tilde{\boldsymbol{w}}$	The rotated moments, defined by (4.32) and (4.36).
ϑ, φ	The Grad-type indices in $\mathcal{D}^m$.
$\varsigma(\cdot)$	Conversion from multi-indices to Grad-type indices.
$\sigma(\cdot)$	Conversion from Grad-type indices to multi-indices.
$\Pi_g(\vartheta, \varphi)$	Product of entries in the rotational matrix, defined by (4.24).
$\mathcal{A}_n^m$	The set of all m-permutations in $\{1, \cdots, n\}$, defined by (4.21).
$\Sigma(\alpha)$	The set of all Grad-type indices whose corresponding multi-indices are α, defined by (4.13).

References

[1] Z. Cai, Y. Fan, and R. Li. Globally hyperbolic regularization of Grad's moment system in one dimensional space. arXiv: 1111.3409, 2012.

[2] Z. Cai and R. Li. Numerical regularized moment method of arbitrary order for Boltzmann-BGK equation. *SIAM J. Sci. Comput.*, 32(5):2875–2907, 2010.

[3] Z. Cai, R. Li, and Z. Qiao. NRxx simulation of microflows with Shakhov model. *SIAM J. Sci. Comput.*, 34(1):A339–A369, 2012.

[4] Z. Cai, R. Li, and Y. Wang. An efficient NRxx method for Boltzmann-BGK equation. *J. Sci. Comput.*, 50(1):103–119, 2012.

[5] Z. Cai, R. Li, and Y. Wang. Numerical regularized moment method for high Mach number flow. *Commun. Comput. Phys.*, 11(5):1415–1438, 2012.

[6] James W. Demmel. *Applied Numerical Linear Algebra*. SIAM, 1997.

[7] H. Grad. On the kinetic theory of rarefied gases. *Comm. Pure Appl. Math.*, 2(4):331–407, 1949.

[8] R. J. Leveque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge, 2002.

[9] C. D. Levermore. Moment closure hierarchies for kinetic theories. *J. Stat. Phys.*, 83(5–6):1021–1065, 1996.

[10] G. Dal Maso, P. G. LeFloch, and F. Murat. Definition and weak stability of nonconservative products. *J. Math. Pures Appl.*, 74(6):483–548, 1995.

[11] I. Müller and T. Ruggeri. *Rational Extended Thermodynamics, Second Edition*, volume 37 of *Springer tracts in natural philosophy*. Springer-Verlag, New York, 1998.

[12] E. F. Toro. *Riemann solvers and numerical methods for fluid dynamics - A practical introduction - 3nd edition*. Springer, 2009.

[13] M. Torrilhon. Special issues on moment methods in kinetic gas theory. *Continuum Mech. Thermodyn.*, 21(5):341–343, 2009.

[14] M. Torrilhon. Hyperbolic moment equations in kinetic gas theory based on multivariate Pearson-IV-distributions. *Commun. Comput. Phys.*, 7(4):639–673, 2010.